(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
26 April 2001 (26.04.2001)

PCT

(10) International Publication Number
WO 01/29176 A2

(51) International Patent Classification[7]: C12N

(21) International Application Number: PCT/US00/28247

(22) International Filing Date: 12 October 2000 (12.10.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/159,860 15 October 1999 (15.10.1999) US

(71) Applicant *(for all designated States except US)*: GENAISSANCE PHARMACEUTICALS, INC. [US/US]; Five Science Park, New Haven, CT 06511 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: CHOI, Julie, Y. [US/US]; 38 Elizabeth Street, West Haven, CT 06516 (US). DENTON, R., Rex [US/US]; 129 Hunters Trail, Madison, CT 06443 (US). NANDABALAN, Krishnan [IN/US]; 228 Village Pond Road, Guilford, CT 06437 (US). STEPHENS, J., Claiborne [US/US]; 46 Crabapple Lane, Guilford, CT 06437 (US).

(74) Agents: SHTIVELBAND, Inna et al.; Genaissance Pharmaceuticals, Inc., Five Science Park, New Haven, CT 06511 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— *Without international search report and to be republished upon receipt of that report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: DRUG TARGET ISOGENES: POLYMORPHISMS IN THE CHOLINERGIC RECEPTOR, MUSCARINIC 3 GENE

(57) Abstract: Polynucleotides comprising one or more of 4 novel single nucleotide polymorphisms in the human cholinergic receptor, muscarinic 3 (CHRM3) gene are described. Compositions and methods for detecting one or more of these polymorphisms are also disclosed. In addition, various genotypes and haplotypes for the CHRM3 gene that exist in the population are described.

WO 01/29176 A2

DRUG TARGET ISOGENES:
POLYMORPHISMS IN THE CHOLINERGIC RECEPTOR, MUSCARINIC 3 GENE

RELATED APPLICATIONS

This application claims the benefit of U.S. Provisional Application Serial No. 60/159,860 filed October 15, 1999.

FIELD OF THE INVENTION

This invention relates to variation in genes that encode pharmaceutically important proteins. In particular, this invention provides genetic variants of the human cholinergic receptor, muscarinic 3 (CHRM3) gene and methods for identifying which variant(s) of this gene is/are possessed by an individual.

BACKGROUND OF THE INVENTION

Current methods for identifying pharmaceuticals to treat disease often start by identifying, cloning, and expressing an important target protein related to the disease. A determination of whether an agonist or antagonist is needed to produce an effect that may benefit a patient with the disease is then made. Then, vast numbers of compounds are screened against the target protein to find new potential drugs. The desired outcome of this process is a drug that is specific for the target, thereby reducing the incidence of the undesired side effects usually caused by a compound's activity at non-intended targets.

What this approach fails to consider, however, is that natural variability exists in any and every population with respect to a particular protein. A target protein currently used to screen drugs typically is expressed by a gene cloned from an individual who was arbitrarily selected. However, the nucleotide sequence of a particular gene may vary tremendously among individuals. Subtle alteration(s) in the primary nucleotide sequence of a gene encoding a target protein may be manifested as significant variation in expression of or in the structure and/or function of the protein. Such alterations may explain the relatively high degree of uncertainty inherent in treatment of individuals with drugs whose design is based upon a single representative example of the target. For example, it is well-established that some classes of drugs frequently have lower efficacy in some individuals than others, which means such individuals and their physicians must weigh the possible benefit of a larger dosage against a greater risk of side effects. In addition, variable information on the biological function or effects of a particular protein may be due to different scientists unknowingly studying different isoforms of the gene encoding the protein. Thus, information on the type and frequency of genomic variation that exists for pharmaceutically important proteins would be useful.

The organization of single nucleotide variations (polymorphisms) in the primary sequence of a gene into one of the limited number of combinations that exist as units of inheritance is termed a

haplotype. Each haplotype therefore contains significantly more information than individual unorganized polymorphisms. Haplotypes provide an accurate measurement of the genomic variation in the two chromosomes of an individual and can be used to identify and distinguish between isoforms of genes coding for drug targets.

It is well-established that many diseases are associated with specific variations in gene sequences. However while there are examples in which individual polymorphisms act as genetic markers for a particular phenotype, in other cases an individual polymorphism may be found in a variety of genomic backgrounds and therefore shows no definitive coupling between the polymorphism and the causative site for the phenotype (Clark AG et al. 1998 *Am J Hum Genet* **63**:595-612; Ulbrecht M et al. 2000 *Am J Respir Crit Care Med* **161**: 469-74). In addition, the marker may be predictive in some populations, but not in other populations (Clark AG et al. 1998 *supra*). In these instances, a haplotype will provide a superior genetic marker for the phenotype (Clark AG et al. 1998 *supra*; Ulbrecht M et al. 2000, *supra*; Ruaño G & Stephens JC *Gen Eng News* **19** (21), December 1999).

Analysis of the association between each observed haplotype and a particular phenotype permits ranking of each haplotype by its statistical power of prediction for the phenotype. Haplotypes found to be strongly associated with the phenotype can then have that positive association confirmed by alternative methods to minimize false positives. For a gene suspected to be associated with a particular phenotype, if no observed haplotypes for that gene show association with the phenotype of interest, then it may be inferred that variation in the gene has little, if any, involvement with that phenotype (Ruaño & Stephens 1999, *supra*). Thus, information on the observed haplotypes and their frequency of occurrence in various population groups will be useful in a variety of research and clinical applications.

One possible drug target for the treatment of Alzheimer's disease, Sjogren's syndrome, disorders associated with smooth muscle contractility and sudden infant death syndrome is the cholinergic receptor, muscarinic 3 (CHRM3) gene or its encoded product. CHRM3, also known as muscarinic cholinergic receptor 3 and HM3, is one subtype of a heterogeneous family of five genetically distinct muscarinic acetylcholine receptors (mAChR) that play important roles in higher brain functions such as learning and memory (Levy, A.I., *Proc. Natl. Acad. Sci.* 93:13541-13546, 1996). Evidence has indicated that impaired cholinergic neurotransmission at mAChRs contributes to the dementia symptoms observed in Alzheimer's disease (Rodriguez-Puertas et al. 1997, Synapse, 26:341-50). CHRM3 plays a key role in salivary secretion, pupillary constriction, and bladder detrusor contractions (Matsui et al. *Proc. Natl. Acad. Sci.* 15; 97(17):9579-84, 2000). Autoantibodies that act as antagonists for CHRM3 on smooth muscle occur in a subset of patients with primary and secondary Sjogren's syndrome, which is an autoimmune disorder characterized by dry eyes and mouth (sicca syndrome) and lymphocytic infiltration of the lacrimal and salivary glands (Waterman et al. *Arthritis Rheum.* 43(7):1647-54, 2000).

CHRM3 antagonists have also been implicated in the treatment of disorders associated with

smooth muscle contractility or tone. These disorders include irritable bowel syndrome, chronic obstructive airway disease, and urinary incontinence (Wallis. Life Sci. 1995;56(11-12):861-8). Specifically, the treatment of obstructive airway disease may be improved by antimuscarinic agents, which selectively block CHRM3, but do not inhibit prejunctional cholinergic autoreceptors that limit release of acetylcholine (Alabaster. Life Sci 1997;60(13-14):1053-60). Furthermore, CHRM3 may also be involved in central respiratory drive and control (Richardson et al. Ann Otol Rhinol Laryngol 1997 Nov;106(11):920-6). Both the levels of CHRM3 messenger RNA in the brain stem and the duration of the laryngeal chemoresponse apnea change according to postnatal age. The laryngeal chemoresponse, comprising laryngeal adductor spasm, central apnea, and subsequent cardiovascular instability, is thought to be a factor in sudden infant death syndrome. Thus, Richardson et al (*supra*) suggest that there is an age-related influence on laryngeal chemoresponse by CHRM3 that decreases with maturation.

The cholinergic receptor, muscarinic 3 gene is located on chromosome 1q41-q44 and contains 1 exon that encodes a 596 amino acid protein. Reference sequences for the CHRM3 gene (GenBank Accession No: U29589.1; SEQ ID NO:1), coding sequence, and protein are shown in Figures 1, 2 and 3, respectively.

Because of the potential for polymorphisms in the CHRM3 gene to affect the expression and function of the encoded protein, it would be useful to determine whether polymorphisms exist in the CHRM3 gene, as well as how such polymorphisms are combined in different copies of the gene. Such information would be useful for studying the biological function of CHRM3 as well as in identifying drugs targeting this protein for the treatment of disorders related to its abnormal expression or function.

SUMMARY OF THE INVENTION

Accordingly, the inventors herein have discovered 4 novel polymorphic sites in the CHRM3 gene. These polymorphic sites (PS) correspond to the following nucleotide positions in the indicated GenBank Accession Number: 53 (PS1), 394 (PS2), 1387 (PS3) and 1493 (PS4) in U29589.1. The polymorphisms at these sites are thymine or cytosine at PS1, guanine or adenine at PS2, cytosine or adenine at PS3 and thymine or cytosine at PS4. In addition, the inventors have determined the identity of the alternative nucleotides present at these sites in a human reference population of 79 unrelated individuals self-identified as belonging to one of four major population groups: African descent, Asian, Caucasian and Hispanic/Latino. It is believed that CHRM3-encoding polynucleotides containing one or more of the novel polymorphic sites reported herein will be useful in studying the expression and biological function of CHRM3, as well as in developing drugs targeting this protein. In addition, information on the combinations of polymorphisms in the CHRM3 gene may have diagnostic and forensic applications.

Thus, in one embodiment, the invention provides an isolated polynucleotide comprising a nucleotide sequence which is a polymorphic variant of a reference sequence for the CHRM3 gene or a

fragment thereof. The reference sequence comprises SEQ ID NO:1 and the polymorphic variant comprises at least one polymorphism selected from the group consisting of cytosine at PS1, adenine at PS2, adenine at PS3 and cytosine at PS4. A particularly preferred polymorphic variant is a naturally-occurring isoform (also referred to herein as an "isogene") of the CHRM3 gene. A CHRM3 isogene of the invention comprises thymine or cytosine at PS1, guanine or adenine at PS2, cytosine or adenine at PS3 and thymine or cytosine at PS4. The invention also provides a collection of CHRM3 isogenes, referred to herein as a CHRM3 genome anthology.

A CHRM3 isogene may be defined by the combination and order of these polymorphisms in the isogene, which is referred to herein as a CHRM3 haplotype. Thus, the invention also provides data on the number of different CHRM3 haplotypes found in the above four population groups. This haplotype data is useful in methods for deriving a CHRM3 haplotype from an individual's genotype for the CHRM3 gene and for determining an association between a CHRM3 haplotype and a particular trait.

In another embodiment, the invention provides a polynucleotide comprising a polymorphic variant of a reference sequence for a CHRM3 cDNA or a fragment thereof. The reference sequence comprises SEQ ID NO:2 (Fig. 2) and the polymorphic cDNA comprises at least one polymorphism selected from the group consisting of adenine at a position corresponding to nucleotide 193, adenine at a position corresponding to nucleotide 1186 and cytosine at a position corresponding to nucleotide 1292.

Polynucleotides complementary to these CHRM3 genomic and cDNA variants are also provided by the invention.

In other embodiments, the invention provides a recombinant expression vector comprising one of the polymorphic genomic variants operably linked to expression regulatory elements as well as a recombinant host cell transformed or transfected with the expression vector. The recombinant vector and host cell may be used to express CHRM3 for protein structure analysis and drug binding studies.

In yet another embodiment, the invention provides a polypeptide comprising a polymorphic variant of a reference amino acid sequence for the CHRM3 protein. The reference amino acid sequence comprises SEQ ID NO:3 (Fig. 3) and the polymorphic variant comprises at least one variant amino acid selected from the group consisting of isoleucine at a position corresponding to amino acid position 65, methionine at a position corresponding to amino acid position 396, and proline at a position corresponding to amino acid position 431. A polymorphic variant of CHRM3 is useful in studying the effect of the variation on the biological activity of CHRM3 as well as on the binding affinity of candidate drugs targeting CHRM3 for the treatment of Alzheimer's disease, Sjogren's syndrome, disorders associated with smooth muscle contractility and sudden infant death syndrome.

The present invention also provides antibodies that recognize and bind to the above polymorphic CHRM3 protein variant. Such antibodies can be utilized in a variety of diagnostic and prognostic formats and therapeutic methods.

In other embodiments, the invention provides methods, compositions, and kits for haplotyping and/or genotyping the CHRM3 gene in an individual. The methods involve identifying the nucleotide or nucleotide pair present at one or more polymorphic sites selected from PS1, PS2, PS3, and PS4 in one or both copies of the CHRM3 gene from the individual. The compositions contain oligonucleotide probes and primers designed to specifically hybridize to one or more target regions containing, or that are adjacent to, a polymorphic site. The methods and compositions for establishing the genotype or haplotype of an individual at the novel polymorphic sites described herein are useful for studying population diversity, anthropological lineage, the significance of diversity and lineage at the phenotypic level, paternity testing, forensic applications, and for identifying associations between the CHRM3 genetic variation and a trait such as level of drug response or susceptibility to disease.

In yet another embodiment, the invention provides a method for identifying an association between a genotype or haplotype and a trait. In preferred embodiments, the trait is susceptibility to a disease, severity of a disease, the staging of a disease or response to a drug. Such methods have applicability in developing diagnostic tests and therapeutic treatments for Alzheimer's disease, Sjogren's syndrome, disorders associated with smooth muscle contractility and sudden infant death syndrome.

The present invention also provides nonhuman transgenic animals comprising one of the CHRM3 genomic polymorphic variants described herein and methods for producing such animals. The transgenic animals are useful for studying expression of the CHRM3 isogenes *in vivo*, for *in vivo* screening and testing of drugs targeted against CHRM3 protein, and for testing the efficacy of therapeutic agents and compounds for Alzheimer's disease, Sjogren's syndrome, disorders associated with smooth muscle contractility and sudden infant death syndrome in a biological system.

The present invention also provides a computer system for storing and displaying polymorphism data determined for the CHRM3 gene. The computer system comprises a computer processing unit; a display; and a database containing the polymorphism data. The polymorphism data includes the polymorphisms, the genotypes and the haplotypes identified for the CHRM3 gene in a reference population. In a preferred embodiment, the computer system is capable of producing a display showing CHRM3 haplotypes organized according to their evolutionary relationships.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 illustrates a reference sequence for the CHRM3 gene (Genbank Version Number U29589.1; contiguous lines; SEQ ID NO:1), with the start and stop positions of each region of coding sequence indicated below the sequence by the numbers within the brackets and the polymorphic sites and polymorphisms identified by Applicants in a reference population indicated by the variant nucleotide positioned below the polymorphic site in the sequence.

Figure 2 illustrates a reference sequence for the CHRM3 coding sequence (contiguous lines; SEQ ID NO:2), with the polymorphic sites and polymorphisms identified by Applicants in a reference

population indicated by the variant nucleotide positioned below the polymorphic site in the sequence.

Figure 3 illustrates a reference sequence for the CHRM3 protein (contiguous lines; SEQ ID NO:3), with the variant amino acids caused by the polymorphisms of Fig. 2 positioned below the polymorphic site in the sequence.

5

DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention is based on the discovery of novel variants of the CHRM3 gene. As described in more detail below, the inventors herein discovered 4 novel polymorphic sites by characterizing the CHRM3 gene found in genomic DNAs isolated from an Index Repository that 10 contains immortalized cell lines from one chimpanzee and 93 human individuals. The human individuals included a reference population of 79 unrelated individuals self-identified as belonging to one of four major population groups: Caucasian (22 individuals), African descent (20 individuals) Asian (20 individuals) Hispanic/Latino (17 individuals). To the extent possible, the members of this reference population were organized into population subgroups by the self-identified ethnogeographic 15 origin of their four grandparents as shown in Table 1 below.

Table 1. Population Groups in the Index Repository

Population Group	Population Subgroup	No. of Individuals
African descent		20
	Sierra Leone	1
Asian		20
	Burma	1
	China	3
	Japan	6
	Korea	1
	Philippines	5
	Vietnam	4
Caucasian		22
	British Isles	3
	British Isles/Central	4
	British Isles/Eastern	1
	Central/Eastern	1
	Eastern	3
	Central/Mediterranean	1
	Mediterranean	2
	Scandinavian	2
Hispanic/Latino		17
	Caribbean	7
	Caribbean (Spanish Descent)	2
	Central American (Spanish Descent)	1
	Mexican American	4
	South American (Spanish Descent)	3

In addition, the Index Repository contains three unrelated indigenous American Indians (one 20 from each of North, Central and South America), one three-generation Caucasian family (from the

CEPH Utah cohort) and one two-generation African-American family.

Using the CHRM3 genotypes identified in the Index Repository and the methodology described in the Examples below, the inventors herein also determined the haplotypes found on each chromosome for most human members of this repository. The CHRM3 genotypes and haplotypes found in the repository include those shown in Tables 4 and 5, respectively. The polymorphism and haplotype data disclosed herein are useful for studying population diversity, anthropological lineage, the significance of diversity and lineage at the phenotypic level, paternity testing, forensic applications, and for identifying associations between the CHRM3 genetic variation and a trait such as level of drug response or susceptibility to disease.

In the context of this disclosure, the following terms shall be defined as follows unless otherwise indicated:

Allele - A particular form of a genetic locus, distinguished from other forms by its particular nucleotide sequence.

Candidate Gene – A gene which is hypothesized to be responsible for a disease, condition, or the response to a treatment, or to be correlated with one of these.

Gene - A segment of DNA that contains all the information for the regulated biosynthesis of an RNA product, including promoters, exons, introns, and other untranslated regions that control expression.

Genotype – An unphased 5′ to 3′ sequence of nucleotide pair(s) found at one or more polymorphic sites in a locus on a pair of homologous chromosomes in an individual. As used herein, genotype includes a full-genotype and/or a sub-genotype as described below.

Full-genotype – The unphased 5′ to 3′ sequence of nucleotide pairs found at all known polymorphic sites in a locus on a pair of homologous chromosomes in a single individual.

Sub-genotype – The unphased 5′ to 3′ sequence of nucleotides seen at a subset of the known polymorphic sites in a locus on a pair of homologous chromosomes in a single individual.

Genotyping – A process for determining a genotype of an individual.

Haplotype – A 5′ to 3′ sequence of nucleotides found at one or more polymorphic sites in a locus on a single chromosome from a single individual. As used herein, haplotype includes a full-haplotype and/or a sub-haplotype as described below.

Full-haplotype – The 5′ to 3′ sequence of nucleotides found at all known polymorphic sites in a locus on a single chromosome from a single individual.

Sub-haplotype – The 5′ to 3′ sequence of nucleotides seen at a subset of the known polymorphic sites in a locus on a single chromosome from a single individual.

Haplotype pair – The two haplotypes found for a locus in a single individual.

Haplotyping – A process for determining one or more haplotypes in an individual and includes use of family pedigrees, molecular techniques and/or statistical inference.

Haplotype data - Information concerning one or more of the following for a specific gene: a

listing of the haplotype pairs in each individual in a population; a listing of the different haplotypes in a population; frequency of each haplotype in that or other populations, and any known associations between one or more haplotypes and a trait.

Isoform – A particular form of a gene, mRNA, cDNA or the protein encoded thereby, distinguished from other forms by its particular sequence and/or structure.

Isogene – One of the isoforms of a gene found in a population. An isogene contains all of the polymorphisms present in the particular isoform of the gene.

Isolated – As applied to a biological molecule such as RNA, DNA, oligonucleotide, or protein, isolated means the molecule is substantially free of other biological molecules such as nucleic acids, proteins, lipids, carbohydrates, or other material such as cellular debris and growth media. Generally, the term "isolated" is not intended to refer to a complete absence of such material or to absence of water, buffers, or salts, unless they are present in amounts that substantially interfere with the methods of the present invention.

Locus - A location on a chromosome or DNA molecule corresponding to a gene or a physical or phenotypic feature.

Naturally-occurring – A term used to designate that the object it is applied to, e.g., naturally-occurring polynucleotide or polypeptide, can be isolated from a source in nature and which has not been intentionally modified by man.

Nucleotide pair – The nucleotides found at a polymorphic site on the two copies of a chromosome from an individual.

Phased – As applied to a sequence of nucleotide pairs for two or more polymorphic sites in a locus, phased means the combination of nucleotides present at those polymorphic sites on a single copy of the locus is known.

Polymorphic site (PS) – A position within a locus at which at least two alternative sequences are found in a population, the most frequent of which has a frequency of no more than 99%.

Polymorphic variant – A gene, mRNA, cDNA, polypeptide or peptide whose nucleotide or amino acid sequence varies from a reference sequence due to the presence of a polymorphism in the gene.

Polymorphism – The sequence variation observed in an individual at a polymorphic site. Polymorphisms include nucleotide substitutions, insertions, deletions and microsatellites and may, but need not, result in detectable differences in gene expression or protein function.

Polymorphism data – Information concerning one or more of the following for a specific gene: location of polymorphic sites; sequence variation at those sites; frequency of polymorphisms in one or more populations; the different genotypes and/or haplotypes determined for the gene; frequency of one or more of these genotypes and/or haplotypes in one or more populations; any known association(s) between a trait and a genotype or a haplotype for the gene.

Polymorphism Database – A collection of polymorphism data arranged in a systematic or

methodical way and capable of being individually accessed by electronic or other means.

Polynucleotide – A nucleic acid molecule comprised of single-stranded RNA or DNA or comprised of complementary, double-stranded DNA.

Population Group – A group of individuals sharing a common ethnogeographic origin.

Reference Population – A group of subjects or individuals who are predicted to be representative of the genetic variation found in the general population. Typically, the reference population represents the genetic variation in the population at a certainty level of at least 85%, preferably at least 90%, more preferably at least 95% and even more preferably at least 99%.

Single Nucleotide Polymorphism (SNP) – Typically, the specific pair of nucleotides observed at a single polymorphic site. In rare cases, three or four nucleotides may be found.

Subject – A human individual whose genotypes or haplotypes or response to treatment or disease state are to be determined.

Treatment - A stimulus administered internally or externally to a subject.

Unphased – As applied to a sequence of nucleotide pairs for two or more polymorphic sites in a locus, unphased means the combination of nucleotides present at those polymorphic sites on a single copy of the locus is not known.

The inventors herein have discovered 4 novel polymorphic sites in the CHRM3 gene. The polymorphic sites identified by the inventors are referred to as PS1-4 to designate the order in which they are located in the gene (see Table 3 below).

Thus, in one embodiment, the invention provides an isolated polynucleotide comprising a polymorphic variant of the CHRM3 gene or a fragment of the gene which contains at least one of the novel polymorphic sites described herein. The nucleotide sequence of a variant CHRM3 gene is identical to the reference genomic sequence for those portions of the gene examined, as described in the Examples below, except that it comprises a different nucleotide at one or more of the novel polymorphic sites PS1, PS2, PS3, and PS4. Similarly, the nucleotide sequence of a variant fragment of the CHRM3 gene is identical to the corresponding portion of the reference sequence except for having a different nucleotide at one or more of the novel polymorphic sites described herein. Thus, the invention specifically does not include polynucleotides comprising a nucleotide sequence identical to the reference sequence (or other reported CHRM3 sequences) or to portions of the reference sequence (or other reported CHRM3 sequences), except for genotyping oligonucleotides as described below.

The location of a polymorphism in a variant gene or fragment is identified by aligning its sequence against SEQ ID NO:1. The polymorphism is selected from the group consisting of cytosine at PS1, adenine at PS2, adenine at PS3 and cytosine at PS4. In a preferred embodiment, the polymorphic variant comprises a naturally-occurring isogene of the CHRM3 gene which is defined by any one of haplotypes 1-3 shown in Table 5 below.

Polymorphic variants of the invention may be prepared by isolating a clone containing the CHRM3 gene from a human genomic library. The clone may be sequenced to determine the identity of

the nucleotides at the polymorphic sites described herein. Any particular variant claimed herein could be prepared from this clone by performing *in vitro* mutagenesis using procedures well-known in the art.

CHRM3 isogenes may be isolated using any method that allows separation of the two "copies" of the CHRM3 gene present in an individual, which, as readily understood by the skilled artisan, may be the same allele or different alleles. Separation methods include targeted *in vivo* cloning (TIVC) in yeast as described in WO 98/01573, U.S. Patent No. 5,866,404, and U.S. Patent No. 5,972,614. Another method, which is described in U.S. Patent No. 5,972,614, uses an allele specific oligonucleotide in combination with primer extension and exonuclease degradation to generate hemizygous DNA targets. Yet other methods are single molecule dilution (SMD) as described in Ruaño et al., Proc. Natl. Acad. Sci. 87:6296-6300, 1990; and allele specific PCR (Ruaño et al., 17 Nucleic Acids. Res. 8392, 1989; Ruaño et al., 19 Nucleic Acids Res. 6877-6882, 1991; Michalatos-Beloin et al., 24 Nucleic Acids Res. 4841-4843, 1996).

The invention also provides CHRM3 genome anthologies, which are collections of CHRM3 isogenes found in a given population. The population may be any group of at least two individuals, including but not limited to a reference population, a population group, a family population, a clinical population, and a same sex population. A CHRM3 genome anthology may comprise individual CHRM3 isogenes stored in separate containers such as microtest tubes, separate wells of a microtitre plate and the like. Alternatively, two or more groups of the CHRM3 isogenes in the anthology may be stored in separate containers. Individual isogenes or groups of isogenes in a genome anthology may be stored in any convenient and stable form, including but not limited to in buffered solutions, as DNA precipitates, freeze-dried preparations and the like. A preferred CHRM3 genome anthology of the invention comprises a set of isogenes defined by the haplotypes shown in Table 5 below.

An isolated polynucleotide containing a polymorphic variant nucleotide sequence of the invention may be operably linked to one or more expression regulatory elements in a recombinant expression vector capable of being propagated and expressing the encoded CHRM3 protein in a prokaryotic or a eukaryotic host cell. Examples of expression regulatory elements which may be used include, but are not limited to, the lac system, operator and promoter regions of phage lambda, yeast promoters, and promoters derived from vaccinia virus, adenovirus, retroviruses, or SV40. Other regulatory elements include, but are not limited to, appropriate leader sequences, termination codons, polyadenylation signals, and other sequences required for the appropriate transcription and subsequent translation of the nucleic acid sequence in a given host cell. Of course, the correct combinations of expression regulatory elements will depend on the host system used. In addition, it is understood that the expression vector contains any additional elements necessary for its transfer to and subsequent replication in the host cell. Examples of such elements include, but are not limited to, origins of replication and selectable markers. Such expression vectors are commercially available or are readily constructed using methods known to those in the art (e.g., F. Ausubel et al., 1987, in "Current Protocols in Molecular Biology", John Wiley and Sons, New York, New York). Host cells which may be used to

express the variant CHRM3 sequences of the invention include, but are not limited to, eukaryotic and mammalian cells, such as animal, plant, insect and yeast cells, and prokaryotic cells, such as E. coli, or algal cells as known in the art. The recombinant expression vector may be introduced into the host cell using any method known to those in the art including, but not limited to, microinjection, electroporation, particle bombardment, transduction, and transfection using DEAE-dextran, lipofection, or calcium phosphate (see e.g., Sambrook et al. (1989) in "Molecular Cloning. A Laboratory Manual", Cold Spring Harbor Press, Plainview, New York). In a preferred aspect, eukaryotic expression vectors that function in eukaryotic cells, and preferably mammalian cells, are used. Non-limiting examples of such vectors include vaccinia virus vectors, adenovirus vectors, herpes virus vectors, and baculovirus transfer vectors. Preferred eukaryotic cell lines include COS cells, CHO cells, HeLa cells, NIH/3T3 cells, and embryonic stem cells (Thomson, J. A. et al., 1998 Science 282:1145-1147). Particularly preferred host cells are mammalian cells.

As will be readily recognized by the skilled artisan, expression of polymorphic variants of the CHRM3 gene will produce CHRM3 mRNAs varying from each other at any polymorphic site retained in the spliced and processed mRNA molecules. These mRNAs can be used for the preparation of a CHRM3 cDNA comprising a nucleotide sequence which is a polymorphic variant of the CHRM3 reference coding sequence shown in Figure 2. Thus, the invention also provides CHRM3 mRNAs and corresponding cDNAs which comprise a nucleotide sequence that is identical to SEQ ID NO:2 (Fig. 2), or its corresponding RNA sequence, except for having one or more polymorphisms selected from the group consisting of adenine at a position corresponding to nucleotide 193, adenine at a position corresponding to nucleotide 1186 and cytosine at a position corresponding to nucleotide 1292. Fragments of these variant mRNAs and cDNAs are included in the scope of the invention, provided they contain the novel polymorphisms described herein. The invention specifically excludes polynucleotides identical to previously identified and characterized CHRM3 cDNAs and fragments thereof. Polynucleotides comprising a variant RNA or DNA sequence may be isolated from a biological sample using well-known molecular biological procedures or may be chemically synthesized.

Polymorphic variants of fragments according to the invention comprise at least one novel polymorphism identified herein and have a length of at least 10 nucleotides and may range up to the full length of the gene. Preferably, such fragments are between 100 and 3000 nucleotides in length, and more preferably between 200 and 2000 nucleotides in length, and most preferably between 500 and 1000 nucleotides in length.

In describing the polymorphic sites identified herein, reference is made to the sense strand of the gene for convenience. However, as recognized by the skilled artisan, nucleic acid molecules containing the CHRM3 gene may be complementary double stranded molecules and thus reference to a particular site on the sense strand refers as well to the corresponding site on the complementary antisense strand. Thus, reference may be made to the same polymorphic site on either strand and an

oligonucleotide may be designed to hybridize specifically to either strand at a target region containing the polymorphic site. Thus, the invention also includes single-stranded polynucleotides which are complementary to the sense strand of the CHRM3 genomic variants described herein.

Polynucleotides comprising a polymorphic gene variant or fragment may be useful for therapeutic purposes. For example, where a patient could benefit from expression, or increased expression, of a particular CHRM3 protein isoform, an expression vector encoding the isoform may be administered to the patient. The patient may be one who lacks the CHRM3 isogene encoding that isoform or may already have at least one copy of that isogene.

In other situations, it may be desirable to decrease or block expression of a particular CHRM3 isogene. Expression of a CHRM3 isogene may be turned off by transforming a targeted organ, tissue or cell population with an expression vector that expresses high levels of untranslatable mRNA for the isogene. Alternatively, oligonucleotides directed against the regulatory regions (e.g., promoter, introns, enhancers, 3′ untranslated region) of the isogene may block transcription. Oligonucleotides targeting the transcription initiation site, e.g., between positions −10 and +10 from the start site are preferred. Similarly, inhibition of transcription can be achieved using oligonucleotides that base-pair with region(s) of the isogene DNA to form triplex DNA (see e.g., Gee et al. in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing Co., Mt. Kisco, N.Y., 1994). Antisense oligonucleotides may also be designed to block translation of CHRM3 mRNA transcribed from a particular isogene. It is also contemplated that ribozymes may be designed that can catalyze the specific cleavage of CHRM3 mRNA transcribed from a particular isogene.

The oligonucleotides may be delivered to a target cell or tissue by expression from a vector introduced into the cell or tissue *in vivo* or *ex vivo*. Alternatively, the oligonucleotides may be formulated as a pharmaceutical composition for administration to the patient. Oligoribonucleotides and/or oligodeoxynucleotides intended for use as antisense oligonucleotides may be modified to increase stability and half-life. Possible modifications include, but are not limited to phosphorothioate or 2′ O-methyl linkages, and the inclusion of nontraditional bases such as inosine and queosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytosine, guanine, thymine, and uracil which are not as easily recognized by endogenous nucleases.

The invention also provides an isolated polypeptide comprising a polymorphic variant of the reference CHRM3 amino acid sequence shown in Figure 3. The location of a variant amino acid in a CHRM3 polypeptide or fragment of the invention is identified by aligning its sequence against SEQ ID NO:3 (Fig. 3). A CHRM3 protein variant of the invention comprises an amino acid sequence identical to SEQ ID NO: 3 except for having one or more variant amino acids selected from the group consisting of isoleucine at a position corresponding to amino acid position 65, methionine at a position corresponding to amino acid position 396, and proline at a position corresponding to amino acid position 431. The invention specifically excludes amino acid sequences identical to those previously identified for CHRM3, including SEQ ID NO: 3, and previously described fragments thereof. CHRM3

protein variants included within the invention comprise all amino acid sequences based on SEQ ID NO: 3 and having the combination of amino acid variations described in Table 2 below. In preferred embodiments, a CHRM3 protein variant of the invention is encoded by an isogene defined by one of the observed haplotypes shown in Table 5.

Table2. Novel Polymorphic Variant of CHRM3

Polymorphic Variant Number	Amino Acid Position and Identities		
	65	396	431
1	I	L	L
2	I	M	P
3	I	M	L
4	I	L	P
5	V	M	P
6	V	M	L
7	V	L	P

The invention also includes CHRM3 peptide variants, which are any fragments of a CHRM3 protein variant that contains one or more of the amino acid variations shown in Table 2. A CHRM3 peptide variant is at least 6 amino acids in length and is preferably any number between 6 and 30 amino acids long, more preferably between 10 and 25, and most preferably between 15 and 20 amino acids long. Such CHRM3 peptide variants may be useful as antigens to generate antibodies specific for one of the above CHRM3 isoforms. In addition, the CHRM3 peptide variants may be useful in drug screening assays.

A CHRM3 variant protein or peptide of the invention may be prepared by chemical synthesis or by expressing one of the variant CHRM3 genomic and cDNA sequences as described above. Alternatively, the CHRM3 protein variant may be isolated from a biological sample of an individual having a CHRM3 isogene which encodes the variant protein. Where the sample contains two different CHRM3 isoforms (i.e., the individual has different CHRM3 isogenes), a particular CHRM3 isoform of the invention can be isolated by immunoaffinity chromatography using an antibody which specifically binds to that particular CHRM3 isoform but does not bind to the other CHRM3 isoform.

The expressed or isolated CHRM3 protein may be detected by methods known in the art, including Coomassie blue staining, silver staining, and Western blot analysis using antibodies specific for the isoform of the CHRM3 protein as discussed further below. CHRM3 variant proteins can be purified by standard protein purification procedures known in the art, including differential precipitation, molecular sieve chromatography, ion-exchange chromatography, isoelectric focusing, gel electrophoresis, affinity and immunoaffinity chromatography and the like. (Ausubel et. al., 1987, In Current Protocols in Molecular Biology John Wiley and Sons, New York, New York). In the case of immunoaffinity chromatography, antibodies specific for a particular polymorphic variant may be used.

A polymorphic variant CHRM3 gene of the invention may also be fused in frame with a heterologous sequence to encode a chimeric CHRM3 protein. The non-CHRM3 portion of the chimeric protein may be recognized by a commercially available antibody. In addition, the chimeric protein may also be engineered to contain a cleavage site located between the CHRM3 and non-

CHRM3 portions so that the CHRM3 protein may be cleaved and purified away from the non-CHRM3 portion.

An additional embodiment of the invention relates to using a novel CHRM3 protein isoform in any of a variety of drug screening assays. Such screening assays may be performed to identify agents that bind specifically to all known CHRM3 protein isoforms or to only a subset of one or more of these isoforms. The agents may be from chemical compound libraries, peptide libraries and the like. The CHRM3 protein or peptide variant may be free in solution or affixed to a solid support. In one embodiment, high throughput screening of compounds for binding to a CHRM3 variant may be accomplished using the method described in PCT application WO84/03565, in which large numbers of test compounds are synthesized on a solid substrate, such as plastic pins or some other surface, contacted with the CHRM3 protein(s) of interest and then washed. Bound CHRM3 protein(s) are then detected using methods well-known in the art.

In another embodiment, a novel CHRM3 protein isoform may be used in assays to measure the binding affinities of one or more candidate drugs targeting the CHRM3 protein.

In another embodiment, the invention provides antibodies specific for and immunoreactive with one or more of the novel CHRM3 variant proteins described herein. The antibodies may be either monoclonal or polyclonal in origin. The CHRM3 protein or peptide variant used to generate the antibodies may be from natural or recombinant sources or produced by chemical synthesis using synthesis techniques known in the art. If the CHRM3 protein variant is of insufficient size to be antigenic, it may be conjugated, complexed, or otherwise covalently linked to a carrier molecule to enhance the antigenicity of the peptide. Examples of carrier molecules, include, but are not limited to, albumins (e.g., human, bovine, fish, ovine), and keyhole limpet hemocyanin (Basic and Clinical Immunology, 1991, Eds. D.P. Stites, and A.I. Terr, Appleton and Lange, Norwalk Connecticut, San Mateo, California).

In one embodiment, an antibody specifically immunoreactive with one of the novel CHRM3 protein isoforms described herein is administered to an individual to neutralize activity of the CHRM3 isoform expressed by that individual. The antibody may be formulated as a pharmaceutical composition which includes a pharmaceutically acceptable carrier.

Antibodies specific for and immunoreactive with one of the novel CHRM3 protein isoform described herein may be used to immunoprecipitate the CHRM3 protein variant from solution as well as react with CHRM3 protein isoforms on Western or immunoblots of polyacrylamide gels on membrane supports or substrates. In another preferred embodiment, the antibodies will detect CHRM3 protein isoforms in paraffin or frozen tissue sections, or in cells which have been fixed or unfixed and prepared on slides, coverslips, or the like, for use in immunocytochemical, immunohistochemical, and immunofluorescence techniques.

In another embodiment, an antibody specifically immunoreactive with one of the novel CHRM3 protein variants described herein is used in immunoassays to detect this variant in biological

samples. In this method, an antibody of the present invention is contacted with a biological sample and the formation of a complex between the CHRM3 protein variant and the antibody is detected. As described, suitable immunoassays include radioimmunoassay, Western blot assay, immunofluorescent assay, enzyme linked immunoassay (ELISA), chemiluminescent assay, immunohistochemical assay, immunocytochemical assay, and the like (see, e.g., Principles and Practice of Immunoassay, 1991, Eds. Christopher P. Price and David J. Neoman, Stockton Press, New York, New York; Current Protocols in Molecular Biology, 1987, Eds. Ausubel et al., John Wiley and Sons, New York, New York). Standard techniques known in the art for ELISA are described in Methods in Immunodiagnosis, 2nd Ed., Eds. Rose and Bigazzi, John Wiley and Sons, New York 1980; and Campbell et al., 1984, Methods in Immunology, W.A. Benjamin, Inc.). Such assays may be direct, indirect, competitive, or noncompetitive as described in the art (see, e.g., Principles and Practice of Immunoassay, 1991, Eds. Christopher P. Price and David J. Neoman, Stockton Pres, NY, NY; and Oellirich, M., 1984, J. Clin. Chem. Clin. Biochem., 22:895-904). Proteins may be isolated from test specimens and biological samples by conventional methods, as described in Current Protocols in Molecular Biology, supra.

Exemplary antibody molecules for use in the detection and therapy methods of the present invention are intact immunoglobulin molecules, substantially intact immunoglobulin molecules, or those portions of immunoglobulin molecules that contain the antigen binding site. Polyclonal or monoclonal antibodies may be produced by methods conventionally known in the art (e.g., Kohler and Milstein, 1975, Nature, 256:495-497; Campbell Monoclonal Antibody Technology, the Production and Characterization of Rodent and Human Hybridomas, 1985, In: Laboratory Techniques in Biochemistry and Molecular Biology, Eds. Burdon et al., Volume 13, Elsevier Science Publishers, Amsterdam). The antibodies or antigen binding fragments thereof may also be produced by genetic engineering. The technology for expression of both heavy and light chain genes in E. coli is the subject of PCT patent applications, publication number WO 901443, WO 901443 and WO 9014424 and in Huse et al., 1989, Science, 246:1275-1281. The antibodies may also be humanized (e.g., Queen, C. et al. 1989 Proc. Natl. Acad. Sci. 86;10029).

Effect(s) of the polymorphisms identified herein on expression of CHRM3 may be investigated by preparing recombinant cells and/or nonhuman recombinant organisms, preferably recombinant animals, containing a polymorphic variant of the CHRM3 gene. As used herein, "expression" includes but is not limited to one or more of the following: transcription of the gene into precursor mRNA; splicing and other processing of the precursor mRNA to produce mature mRNA; mRNA stability; translation of the mature mRNA into CHRM3 protein (including codon usage and tRNA availability); and glycosylation and/or other modifications of the translation product, if required for proper expression and function.

To prepare a recombinant cell of the invention, the desired CHRM3 isogene may be introduced into the cell in a vector such that the isogene remains extrachromosomal. In such a situation, the gene will be expressed by the cell from the extrachromosomal location. In a preferred embodiment, the

CHRM3 isogene is introduced into a cell in such a way that it recombines with the endogenous CHRM3 gene present in the cell. Such recombination requires the occurrence of a double recombination event, thereby resulting in the desired CHRM3 gene polymorphism. Vectors for the introduction of genes both for recombination and for extrachromosomal maintenance are known in the art, and any suitable vector or vector construct may be used in the invention. Methods such as electroporation, particle bombardment, calcium phosphate co-precipitation and viral transduction for introducing DNA into cells are known in the art; therefore, the choice of method may lie with the competence and preference of the skilled practitioner. Examples of cells into which the CHRM3 isogene may be introduced include, but are not limited to, continuous culture cells, such as COS, NIH/3T3, and primary or culture cells of the relevant tissue type, i.e., they express the CHRM3 isogene. Such recombinant cells can be used to compare the biological activities of the different protein variants.

Recombinant nonhuman organisms, i.e., transgenic animals, expressing a variant CHRM3 gene are prepared using standard procedures known in the art. Preferably, a construct comprising the variant gene is introduced into a nonhuman animal or an ancestor of the animal at an embryonic stage, i.e., the one-cell stage, or generally not later than about the eight-cell stage. Transgenic animals carrying the constructs of the invention can be made by several methods known to those having skill in the art. One method involves transfecting into the embryo a retrovirus constructed to contain one or more insulator elements, a gene or genes of interest, and other components known to those skilled in the art to provide a complete shuttle vector harboring the insulated gene(s) as a transgene, see e.g., U.S. Patent No. 5,610,053. Another method involves directly injecting a transgene into the embryo. A third method involves the use of embryonic stem cells. Examples of animals into which the CHRM3 isogenes may be introduced include, but are not limited to, mice, rats, other rodents, and nonhuman primates (see "The Introduction of Foreign Genes into Mice" and the cited references therein, In: Recombinant DNA, Eds. J.D. Watson, M. Gilman, J. Witkowski, and M. Zoller; W.H. Freeman and Company, New York, pages 254-272). Transgenic animals stably expressing a human CHRM3 isogene and producing human CHRM3 protein can be used as biological models for studying diseases related to abnormal CHRM3 expression and/or activity, and for screening and assaying various candidate drugs, compounds, and treatment regimens to reduce the symptoms or effects of these diseases.

An additional embodiment of the invention relates to pharmaceutical compositions for treating disorders affected by expression or function of a novel CHRM3 isogene described herein. The pharmaceutical composition may comprise any of the following active ingredients: a polynucleotide comprising one of these novel CHRM3 isogenes; an antisense oligonucleotide directed against one of the novel CHRM3 isogenes, a polynucleotide encoding such an antisense oligonucleotide, or another compound which inhibits expression of a novel CHRM3 isogene described herein. Preferably, the composition contains the active ingredient in a therapeutically effective amount. By therapeutically effective amount is meant that one or more of the symptoms relating to disorders affected by

expression or function of a novel CHRM3 isogene is reduced and/or eliminated. The composition also comprises a pharmaceutically acceptable carrier, examples of which include, but are not limited to, saline, buffered saline, dextrose, and water. Those skilled in the art may employ a formulation most suitable for the active ingredient, whether it is a polynucleotide, oligonucleotide, protein, peptide or small molecule antagonist. The pharmaceutical composition may be administered alone or in combination with at least one other agent, such as a stabilizing compound. Administration of the pharmaceutical composition may be by any number of routes including, but not limited to oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, intradermal, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing Co., Easton, PA).

For any composition, determination of the therapeutically effective dose of active ingredient and/or the appropriate route of administration is well within the capability of those skilled in the art. For example, the dose can be estimated initially either in cell culture assays or in animal models. The animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans. The exact dosage will be determined by the practitioner, in light of factors relating to the patient requiring treatment, including but not limited to severity of the disease state, general health, age, weight and gender of the patient, diet, time and frequency of administration, other drugs being taken by the patient, and tolerance/response to the treatment.

Information on the identity of genotypes and haplotypes for the CHRM3 gene of any particular individual as well as the frequency of such genotypes and haplotypes in any particular population of individuals is expected to be useful for a variety of basic research and clinical applications. Thus, the invention also provides compositions and methods for detecting the novel CHRM3 polymorphisms identified herein.

The compositions comprise at least one CHRM3 genotyping oligonucleotide. In one embodiment, a CHRM3 genotyping oligonucleotide is a probe or primer capable of hybridizing to a target region that is located close to, or that contains, one of the novel polymorphic sites described herein. As used herein, the term "oligonucleotide" refers to a polynucleotide molecule having less than about 100 nucleotides. A preferred oligonucleotide of the invention is 10 to 35 nucleotides long. More preferably, the oligonucleotide is between 15 and 30, and most preferably, between 20 and 25 nucleotides in length. The oligonucleotide may be comprised of any phosphorylation state of ribonucleotides, deoxyribonucleotides, and acyclic nucleotide derivatives, and other functionally equivalent derivatives. Alternatively, oligonucleotides may have a phosphate-free backbone, which may be comprised of linkages such as carboxymethyl, acetamidate, carbamate, polyamide (peptide nucleic acid (PNA)) and the like (Varma, R. in Molecular Biology and Biotechnology, A Comprehensive Desk Reference, Ed. R. Meyers, VCH Publishers, Inc. (1995), pages 617-620).

Oligonucleotides of the invention may be prepared by chemical synthesis using any suitable methodology known in the art, or may be derived from a biological sample, for example, by restriction digestion. The oligonucleotides may be labeled, according to any technique known in the art, including use of radiolabels, fluorescent labels, enzymatic labels, proteins, haptens, antibodies, sequence tags and the like.

Genotyping oligonucleotides of the invention must be capable of specifically hybridizing to a target region of a CHRM3 polynucleotide, i.e., a CHRM3 isogene. As used herein, specific hybridization means the oligonucleotide forms an anti-parallel double-stranded structure with the target region under certain hybridizing conditions, while failing to form such a structure when incubated with a non-target region or a non-CHRM3 polynucleotide under the same hybridizing conditions. Preferably, the oligonucleotide specifically hybridizes to the target region under conventional high stringency conditions. The skilled artisan can readily design and test oligonucleotide probes and primers suitable for detecting polymorphisms in the CHRM3 gene using the polymorphism information provided herein in conjunction with the known sequence information for the CHRM3 gene and routine techniques.

A nucleic acid molecule such as an oligonucleotide or polynucleotide is said to be a "perfect" or "complete" complement of another nucleic acid molecule if every nucleotide of one of the molecules is complementary to the nucleotide at the corresponding position of the other molecule. A nucleic acid molecule is "substantially complementary" to another molecule if it hybridizes to that molecule with sufficient stability to remain in a duplex form under conventional low-stringency conditions. Conventional hybridization conditions are described, for example, by Sambrook J. et al., in Molecular Cloning, A Laboratory Manual, 2nd Edition, Cold Spring Harbor Press, Cold Spring Harbor, NY (1989) and by Haymes, B.D. et al. in Nucleic Acid Hybridization, A Practical Approach, IRL Press, Washington, D.C. (1985). While perfectly complementary oligonucleotides are preferred for detecting polymorphisms, departures from complete complementarity are contemplated where such departures do not prevent the molecule from specifically hybridizing to the target region. For example, an oligonucleotide primer may have a non-complementary fragment at its 5′ end, with the remainder of the primer being complementary to the target region. Alternatively, non-complementary nucleotides may be interspersed into the oligonucleotide probe or primer as long as the resulting probe or primer is still capable of specifically hybridizing to the target region.

Preferred genotyping oligonucleotides of the invention are allele-specific oligonucleotides. As used herein, the term allele-specific oligonucleotide (ASO) means an oligonucleotide that is able, under sufficiently stringent conditions, to hybridize specifically to one allele of a gene, or other locus, at a target region containing a polymorphic site while not hybridizing to the corresponding region in another allele(s). As understood by the skilled artisan, allele-specificity will depend upon a variety of readily optimized stringency conditions, including salt and formamide concentrations, as well as temperatures for both the hybridization and washing steps. Examples of hybridization and washing

conditions typically used for ASO probes are found in Kogan et al., "Genetic Prediction of Hemophilia A" in PCR Protocols, A Guide to Methods and Applications, Academic Press, 1990 and Ruano et al., 87 Proc. Natl. Acad. Sci. USA 6296-6300, 1990. Typically, an allele-specific oligonucleotide will be perfectly complementary to one allele while containing a single mismatch for another allele.

Allele-specific oligonucleotide probes which usually provide good discrimination between different alleles are those in which a central position of the oligonucleotide probe aligns with the polymorphic site in the target region (e.g., approximately the 7th or 8th position in a 15 mer, the 8th or 9th position in a 16mer, the 10th or 11th position in a 20 mer). A preferred ASO probe for detecting CHRM3 gene polymorphisms comprises a nucleotide sequence, listed 5′ to 3′, selected from the group consisting of:

Accession No.: U29589.1

ATGAACTTGATGTTT (SEQ ID NO:4) and its complement,
ATGAACTCGATGTTT (SEQ ID NO:5) and its complement,
TCATACCGTCTGGCA (SEQ ID NO:6) and its complement,
TCATACCATCTGGCA (SEQ ID NO:7) and its complement,
GGAGGAGCTGGGGAT (SEQ ID NO:8) and its complement,
GGAGGAGATGGGGAT (SEQ ID NO:9) and its complement,
ATCCAGCTAGAGTCA (SEQ ID NO:10) and its complement, and
ATCCAGCCAGAGTCA (SEQ ID NO:11) and its complement.

An allele-specific oligonucleotide primer of the invention has a 3′ terminal nucleotide, or preferably a 3′ penultimate nucleotide, that is complementary to only one nucleotide of a particular SNP, thereby acting as a primer for polymerase-mediated extension only if the allele containing that nucleotide is present. Allele-specific oligonucleotide primers hybridizing to either the coding or noncoding strand are contemplated by the invention. A preferred ASO primer for detecting CHRM3 gene polymorphisms comprises a nucleotide sequence, listed 5′ to 3′, selected from the group consisting of:

Accession No.:U29589.1

AAGAGAATGAACTTG (SEQ ID NO:12); GAAGCCAAACATCAA (SEQ ID NO:13);
AAGAGAATGAACTCG (SEQ ID NO:14); GAAGCCAAACATCGA (SEQ ID NO:15);
GGGAGGTCATACCGT (SEQ ID NO:16); ACCACTTGCCAGACG (SEQ ID NO:17);
GGGAGGTCATACCAT (SEQ ID NO:18); ACCACTTGCCAGATG (SEQ ID NO:19);
GCCTGAGGAGGAGCT (SEQ ID NO:20); TCCACCATCCCCAGC (SEQ ID NO:21);
GCCTGAGGAGGAGAT (SEQ ID NO:22); TCCACCATCCCCATC (SEQ ID NO:23);
CTTCCCATCCAGCTA (SEQ ID NO:24); CACGGCTGACTCTAG (SEQ ID NO:25);
CTTCCCATCCAGCCA (SEQ ID NO:26); and CACGGCTGACTCTGG (SEQ ID NO:27).

Other genotyping oligonucleotides of the invention hybridize to a target region located one to several nucleotides downstream of one of the novel polymorphic sites identified herein. Such oligonucleotides are useful in polymerase-mediated primer extension methods for detecting one of the

novel polymorphisms described herein and therefore such genotyping oligonucleotides are referred to herein as "primer-extension oligonucleotides". In a preferred embodiment, the 3'-terminus of a primer-extension oligonucleotide is a deoxynucleotide complementary to the nucleotide located immediately adjacent to the polymorphic site. A particularly preferred oligonucleotide primer for detecting CHRM3 gene polymorphisms by primer extension terminates in a nucleotide sequence, listed 5' to 3', selected from the group consisting of:

Accession No.:U29589.1

AGAATGAACT	(SEQ ID NO:28);	GCCAAACATC	(SEQ ID NO:29);
AGGTCATACC	(SEQ ID NO:30);	ACTTGCCAGA	(SEQ ID NO:31);
TGAGGAGGAG	(SEQ ID NO:32);	ACCATCCCCA	(SEQ ID NO:33);
CCCATCCAGC	(SEQ ID NO:34); and	GGCTGACTCT	(SEQ ID NO:35).

In some embodiments, a composition contains two or more differently labeled genotyping oligonucleotides for simultaneously probing the identity of nucleotides at two or more polymorphic sites. It is also contemplated that primer compositions may contain two or more sets of allele-specific primer pairs to allow simultaneous targeting and amplification of two or more regions containing a polymorphic site.

CHRM3 genotyping oligonucleotides of the invention may also be immobilized on or synthesized on a solid surface such as a microchip, bead, or glass slide (see, e.g., WO 98/20020 and WO 98/20019). Such immobilized genotyping oligonucleotides may be used in a variety of polymorphism detection assays, including but not limited to probe hybridization and polymerase extension assays. Immobilized CHRM3 genotyping oligonucleotides of the invention may comprise an ordered array of oligonucleotides designed to rapidly screen a DNA sample for polymorphisms in multiple genes at the same time.

In another embodiment, the invention provides a kit comprising at least two genotyping oligonucleotides packaged in separate containers. The kit may also contain other components such as hybridization buffer (where the oligonucleotides are to be used as a probe) packaged in a separate container. Alternatively, where the oligonucleotides are to be used to amplify a target region, the kit may contain, packaged in separate containers, a polymerase and a reaction buffer optimized for primer extension mediated by the polymerase, such as PCR.

The above described oligonucleotide compositions and kits are useful in methods for genotyping and/or haplotyping the CHRM3 gene in an individual. As used herein, the terms "CHRM3 genotype" and "CHRM3 haplotype" mean the genotype or haplotype contains the nucleotide pair or nucleotide, respectively, that is present at one or more of the novel polymorphic sites described herein and may optionally also include the nucleotide pair or nucleotide present at one or more additional polymorphic sites in the CHRM3 gene. The additional polymorphic sites may be currently known polymorphic sites or sites that are subsequently discovered.

One embodiment of the genotyping method involves isolating from the individual a nucleic

acid mixture comprising the two copies of the CHRM3 gene, or a fragment thereof, that are present in the individual, and determining the identity of the nucleotide pair at one or more of the polymorphic sites selected from PS1, PS2, PS3, and PS4 in the two copies to assign a CHRM3 genotype to the individual. As will be readily understood by the skilled artisan, the two "copies" of a gene in an individual may be the same allele or may be different alleles. In a particularly preferred embodiment, the genotyping method comprises determining the identity of the nucleotide pair at each of PS1-4.

Typically, the nucleic acid mixture is isolated from a biological sample taken from the individual, such as a blood sample or tissue sample. Suitable tissue samples include whole blood, semen, saliva, tears, urine, fecal material, sweat, buccal, skin and hair. The nucleic acid mixture may be comprised of genomic DNA, mRNA, or cDNA and, in the latter two cases, the biological sample must be obtained from an organ in which the CHRM3 gene is expressed. Furthermore it will be understood by the skilled artisan that mRNA or cDNA preparations would not be used to detect polymorphisms located in introns or in 5′ and 3′ nontranscribed regions. If a CHRM3 gene fragment is isolated, it must contain the polymorphic site(s) to be genotyped.

One embodiment of the haplotyping method comprises isolating from the individual a nucleic acid molecule containing only one of the two copies of the CHRM3 gene, or a fragment thereof, that is present in the individual and determining in that copy the identity of the nucleotide at one or more of the polymorphic sites PS1, PS2, PS3, and PS4 in that copy to assign a CHRM3 haplotype to the individual. The nucleic acid may be isolated using any method capable of separating the two copies of the CHRM3 gene or fragment such as one of the methods described above for preparing CHRM3 isogenes, with targeted *in vivo* cloning being the preferred approach. As will be readily appreciated by those skilled in the art, any individual clone will only provide haplotype information on one of the two CHRM3 gene copies present in an individual. If haplotype information is desired for the individual's other copy, additional CHRM3 clones will need to be examined. Typically, at least five clones should be examined to have more than a 90% probability of haplotyping both copies of the CHRM3 gene in an individual. In a particularly preferred embodiment, the nucleotide at each of PS1-4 is identified.

In a preferred embodiment, a CHRM3 haplotype pair is determined for an individual by identifying the phased sequence of nucleotides at one or more of the polymorphic sites selected from PS1, PS2, PS3, and PS4 in each copy of the CHRM3 gene that is present in the individual. In a particularly preferred embodiment, the haplotyping method comprises identifying the phased sequence of nucleotides at each of PS1-4 in each copy of the CHRM3 gene. When haplotyping both copies of the gene, the identifying step is preferably performed with each copy of the gene being placed in separate containers. However, it is also envisioned that if the two copies are labeled with different tags, or are otherwise separately distinguishable or identifiable, it could be possible in some cases to perform the method in the same container. For example, if first and second copies of the gene are labeled with different first and second fluorescent dyes, respectively, and an allele-specific oligonucleotide labeled with yet a third different fluorescent dye is used to assay the polymorphic

site(s), then detecting a combination of the first and third dyes would identify the polymorphism in the first gene copy while detecting a combination of the second and third dyes would identify the polymorphism in the second gene copy.

In both the genotyping and haplotyping methods, the identity of a nucleotide (or nucleotide pair) at a polymorphic site(s) may be determined by amplifying a target region(s) containing the polymorphic site(s) directly from one or both copies of the CHRM3 gene, or fragment thereof, and the sequence of the amplified region(s) determined by conventional methods. It will be readily appreciated by the skilled artisan that only one nucleotide will be detected at a polymorphic site in individuals who are homozygous at that site, while two different nucleotides will be detected if the individual is heterozygous for that site. The polymorphism may be identified directly, known as positive-type identification, or by inference, referred to as negative-type identification. For example, where a SNP is known to be guanine and cytosine in a reference population, a site may be positively determined to be either guanine or cytosine for an individual homozygous at that site, or both guanine and cytosine, if the individual is heterozygous at that site. Alternatively, the site may be negatively determined to be not guanine (and thus cytosine/cytosine) or not cytosine (and thus guanine/guanine).

In addition, the identity of the allele(s) present at any of the novel polymorphic sites described herein may be indirectly determined by genotyping a polymorphic site not disclosed herein that is in linkage disequilibrium with the polymorphic site that is of interest. Two sites are said to be in linkage disequilibrium if the presence of a particular variant at one site enhances the predictability of another variant at the second site (Stevens, JC 1999, *Mol. Diag.* 4: 309-17). Polymorphic sites in linkage disequilibrium with the presently disclosed polymorphic sites may be located in regions of the gene or in other genomic regions not examined herein. Genotyping of a polymorphic site in linkage disequilibrium with the novel polymorphic sites described herein may be performed by, but is not limited to, any of the above-mentioned methods for detecting the identity of the allele at a polymorphic site.

The target region(s) may be amplified using any oligonucleotide-directed amplification method, including but not limited to polymerase chain reaction (PCR) (U.S. Patent No. 4,965,188), ligase chain reaction (LCR) (Barany et al., Proc. Natl. Acad. Sci. USA 88:189-193, 1991; WO90/01069), and oligonucleotide ligation assay (OLA) (Landegren et al., Science 241:1077-1080, 1988).

Oligonucleotides useful as primers or probes in such methods should specifically hybridize to a region of the nucleic acid that contains or is adjacent to the polymorphic site. Typically, the oligonucleotides are between 10 and 35 nucleotides in length and preferably, between 15 and 30 nucleotides in length. Most preferably, the oligonucleotides are 20 to 25 nucleotides long. The exact length of the oligonucleotide will depend on many factors that are routinely considered and practiced by the skilled artisan.

Other known nucleic acid amplification procedures may be used to amplify the target region including transcription-based amplification systems (U.S. Patent No. 5,130,238; EP 329,822; U.S.

Patent No. 5,169,766, WO89/06700) and isothermal methods (Walker et al., *Proc. Natl. Acad. Sci. USA* 89:392-396, 1992).

A polymorphism in the target region may also be assayed before or after amplification using one of several hybridization-based methods known in the art. Typically, allele-specific oligonucleotides are utilized in performing such methods. The allele-specific oligonucleotides may be used as differently labeled probe pairs, with one member of the pair showing a perfect match to one variant of a target sequence and the other member showing a perfect match to a different variant. In some embodiments, more than one polymorphic site may be detected at once using a set of allele-specific oligonucleotides or oligonucleotide pairs. Preferably, the members of the set have melting temperatures within 5°C, and more preferably within 2°C, of each other when hybridizing to each of the polymorphic sites being detected.

Hybridization of an allele-specific oligonucleotide to a target polynucleotide may be performed with both entities in solution, or such hybridization may be performed when either the oligonucleotide or the target polynucleotide is covalently or noncovalently affixed to a solid support. Attachment may be mediated, for example, by antibody-antigen interactions, poly-L-Lys, streptavidin or avidin-biotin, salt bridges, hydrophobic interactions, chemical linkages, UV cross-linking baking, etc. Allele-specific oligonucleotides may be synthesized directly on the solid support or attached to the solid support subsequent to synthesis. Solid-supports suitable for use in detection methods of the invention include substrates made of silicon, glass, plastic, paper and the like, which may be formed, for example, into wells (as in 96-well plates), slides, sheets, membranes, fibers, chips, dishes, and beads. The solid support may be treated, coated or derivatized to facilitate the immobilization of the allele-specific oligonucleotide or target nucleic acid.

The genotype or haplotype for the CHRM3 gene of an individual may also be determined by hybridization of a nucleic acid sample containing one or both copies of the gene to nucleic acid arrays and subarrays such as described in WO 95/11995. The arrays would contain a battery of allele-specific oligonucleotides representing each of the polymorphic sites to be included in the genotype or haplotype.

The identity of polymorphisms may also be determined using a mismatch detection technique, including but not limited to the RNase protection method using riboprobes (Winter et al., Proc. Natl. Acad. Sci. USA 82:7575, 1985; Meyers et al., Science 230:1242, 1985) and proteins which recognize nucleotide mismatches, such as the E. coli mutS protein (Modrich, P. Ann. Rev. Genet. 25:229-253, 1991). Alternatively, variant alleles can be identified by single strand conformation polymorphism (SSCP) analysis (Orita et al., Genomics 5:874-879, 1989; Humphries et al., in Molecular Diagnosis of Genetic Diseases, R. Elles, ed., pp. 321-340, 1996) or denaturing gradient gel electrophoresis (DGGE) (Wartell et al., Nucl. Acids Res. 18:2699-2706, 1990; Sheffield et al., Proc. Natl. Acad. Sci. USA 86:232-236, 1989).

A polymerase-mediated primer extension method may also be used to identify the

polymorphism(s). Several such methods have been described in the patent and scientific literature and include the "Genetic Bit Analysis" method (WO92/15712) and the ligase/polymerase mediated genetic bit analysis (U.S. Patent 5,679,524. Related methods are disclosed in WO91/02087, WO90/09455, WO95/17676, U.S. Patent Nos. 5,302,509, and 5,945,283. Extended primers containing a polymorphism may be detected by mass spectrometry as described in U.S. Patent No. 5,605,798. Another primer extension method is allele-specific PCR (Ruaño et al., Nucl. Acids Res. 17:8392, 1989; Ruaño et al., Nucl. Acids Res. 19, 6877-6882, 1991; WO 93/22456; Turki et al., J. Clin. Invest. 95:1635-1641, 1995). In addition, multiple polymorphic sites may be investigated by simultaneously amplifying multiple regions of the nucleic acid using sets of allele-specific primers as described in Wallace et al. (WO89/10414).

In another aspect of the invention, an individual's CHRM3 haplotype pair is predicted from its CHRM3 genotype using information on haplotype pairs known to exist in a reference population. In its broadest embodiment, the haplotyping prediction method comprises identifying a CHRM3 genotype for the individual at two or more polymorphic sites selected from PS1, PS2, PS3, and PS4, enumerating all possible haplotype pairs which are consistent with the genotype, accessing data containing CHRM3 haplotype pairs identified in a reference population, and assigning a haplotype pair to the individual that is consistent with the data. In one embodiment, the reference haplotype pairs include the CHRM3 haplotype pairs shown in Table 4.

Generally, the reference population should be composed of randomly-selected individuals representing the major ethnogeographic groups of the world. A preferred reference population for use in the methods of the present invention comprises an approximately equal number of individuals from Caucasian, African American, Asian and Hispanic-Latino population groups with the minimum number of each group being chosen based on how rare a haplotype one wants to be guaranteed to see. For example, if one wants to have a q% chance of not missing a haplotype that exists in the population at a p% frequency of occurring in the reference population, the number of individuals (n) who must be sampled is given by $2n=\log(1-q)/\log(1-p)$ where p and q are expressed as fractions. A preferred reference population allows the detection of any haplotype whose frequency is at least 10% with about 99% certainty and comprises about 20 unrelated individuals from each of the four population groups named above. A particularly preferred reference population includes a 3-generation family representing one or more of the four population groups to serve as controls for checking quality of haplotyping procedures.

In a preferred embodiment, the haplotype frequency data for each ethnogeographic group is examined to determine whether it is consistent with Hardy-Weinberg equilibrium. Hardy-Weinberg equilibrium (D.L. Hartl et al., Principles of Population Genomics, Sinauer Associates (Sunderland, MA), 3rd Ed., 1997) postulates that the frequency of finding the haplotype pair H_1 / H_2 is equal to $p_{H\text{-}W}(H_1 / H_2) = 2p(H_1)p(H_2)$ if $H_1 \neq H_2$ and $p_{H\text{-}W}(H_1 / H_2) = p(H_1)p(H_2)$ if $H_1 = H_2$. A statistically significant difference between the observed and expected haplotype frequencies could be

due to one or more factors including significant inbreeding in the population group, strong selective pressure on the gene, sampling bias, and/or errors in the genotyping process. If large deviations from Hardy-Weinberg equilibrium are observed in an ethnogeographic group, the number of individuals in that group can be increased to see if the deviation is due to a sampling bias. If a larger sample size does not reduce the difference between observed and expected haplotype pair frequencies, then one may wish to consider haplotyping the individual using a direct haplotyping method such as, for example, CLASPER System™ technology (U.S. Patent No. 5,866,404), SMD, or allele-specific long-range PCR (Michalotos-Beloin et al., Nucleic Acids Res. 24:4841-4843, 1996).

In one embodiment of this method for predicting a CHRM3 haplotype pair, the assigning step involves performing the following analysis. First, each of the possible haplotype pairs is compared to the haplotype pairs in the reference population. Generally, only one of the haplotype pairs in the reference population matches a possible haplotype pair and that pair is assigned to the individual. Occasionally, only one haplotype represented in the reference haplotype pairs is consistent with a possible haplotype pair for an individual, and in such cases the individual is assigned a haplotype pair containing this known haplotype and a new haplotype derived by subtracting the known haplotype from the possible haplotype pair. In rare cases, either no haplotypes in the reference population are consistent with the possible haplotype pairs, or alternatively, multiple reference haplotype pairs are consistent with the possible haplotype pairs. In such cases, the individual is preferably haplotyped using a direct molecular haplotyping method such as, for example, CLASPER System™ technology (U.S. Patent No. 5,866,404), SMD, or allele-specific long-range PCR (Michalotos-Beloin et al., Nucleic Acids Res. 24:4841-4843, 1996). A preferred process for predicting CHRM3 haplotype pairs from CHRM3 genotypes is described in copending U.S. Provisional Application Serial No. 60/198,340.

The invention also provides a method for determining the frequency of a CHRM3 genotype or CHRM3 haplotype in a population. The method comprises determining the genotype or the haplotype pair for the CHRM3 gene that is present in each member of the population, wherein the genotype or haplotype comprises the nucleotide pair or nucleotide detected at one or more of the polymorphic sites PS1, PS2, PS3, and PS4 in the CHRM3 gene; and calculating the frequency any particular genotype or haplotype is found in the population. The population may be a reference population, a family population, a same sex population, a population group, a trait population (e.g., a group of individuals exhibiting a trait of interest such as a medical condition or response to a therapeutic treatment).

In another aspect of the invention, frequency data for CHRM3 genotypes and/or haplotypes found in a reference population are used in a method for identifying an association between a trait and a CHRM3 genotype or a CHRM3 haplotype. The trait may be any detectable phenotype, including but not limited to susceptibility to a disease or response to a treatment. The method involves obtaining data on the frequency of the genotype(s) or haplotype(s) of interest in a reference population as well as in a population exhibiting the trait. Frequency data for one or both of the reference and trait

populations may be obtained by genotyping or haplotyping each individual in the populations using one of the methods described above. The haplotypes for the trait population may be determined directly or, alternatively, by the predictive genotype to haplotype approach described above. In another embodiment, the frequency data for the reference and/or trait populations is obtained by accessing previously determined frequency data, which may be in written or electronic form. For example, the frequency data may be present in a database that is accessible by a computer. Once the frequency data is obtained, the frequencies of the genotype(s) or haplotype(s) of interest in the reference and trait populations are compared. In a preferred embodiment, the frequencies of all genotypes and/or haplotypes observed in the populations are compared. If a particular genotype or haplotype for the CHRM3 gene is more frequent in the trait population than in the reference population at a statistically significant amount, then the trait is predicted to be associated with that CHRM3 genotype or haplotype. Preferably, the CHRM3 genotype or haplotype being compared in the trait and reference populations is selected from the full-genotypes and full-haplotypes shown in Tables 4 and 5, respectively, or from sub-genotypes and sub-haplotypes derived from these genotypes and haplotypes.

In a preferred embodiment of the method, the trait of interest is a clinical response exhibited by a patient to some therapeutic treatment, for example, response to a drug targeting CHRM3 or response to a therapeutic treatment for a medical condition. As used herein, "medical condition" includes but is not limited to any condition or disease manifested as one or more physical and/or psychological symptoms for which treatment is desirable, and includes previously and newly identified diseases and other disorders. As used herein the term "clinical response" means any or all of the following: a quantitative measure of the response, no response, and adverse response (i.e., side effects).

In order to deduce a correlation between clinical response to a treatment and a CHRM3 genotype or haplotype, it is necessary to obtain data on the clinical responses exhibited by a population of individuals who received the treatment, hereinafter the "clinical population". This clinical data may be obtained by analyzing the results of a clinical trial that has already been run and/or the clinical data may be obtained by designing and carrying out one or more new clinical trials. As used herein, the term "clinical trial" means any research study designed to collect clinical data on responses to a particular treatment, and includes but is not limited to phase I, phase II and phase III clinical trials. Standard methods are used to define the patient population and to enroll subjects.

It is preferred that the individuals included in the clinical population have been graded for the existence of the medical condition of interest. This is important in cases where the symptom(s) being presented by the patients can be caused by more than one underlying condition, and where treatment of the underlying conditions are not the same. An example of this would be where patients experience breathing difficulties that are due to either asthma or respiratory infections. If both sets were treated with an asthma medication, there would be a spurious group of apparent non-responders that did not actually have asthma. These people would affect the ability to detect any correlation between haplotype and treatment outcome. This grading of potential patients could employ a standard physical

exam or one or more lab tests. Alternatively, grading of patients could use haplotyping for situations where there is a strong correlation between haplotype pair and disease susceptibility or severity.

The therapeutic treatment of interest is administered to each individual in the trial population and each individual's response to the treatment is measured using one or more predetermined criteria. It is contemplated that in many cases, the trial population will exhibit a range of responses and that the investigator will choose the number of responder groups (e.g., low, medium, high) made up by the various responses. In addition, the CHRM3 gene for each individual in the trial population is genotyped and/or haplotyped, which may be done before or after administering the treatment.

After both the clinical and polymorphism data have been obtained, correlations between individual response and CHRM3 genotype or haplotype content are created. Correlations may be produced in several ways. In one method, individuals are grouped by their CHRM3 genotype or haplotype (or haplotype pair) (also referred to as a polymorphism group), and then the averages and standard deviations of clinical responses exhibited by the members of each polymorphism group are calculated.

These results are then analyzed to determine if any observed variation in clinical response between polymorphism groups is statistically significant. Statistical analysis methods which may be used are described in L.D. Fisher and G. vanBelle, "Biostatistics: A Methodology for the Health Sciences", Wiley-Interscience (New York) 1993. This analysis may also include a regression calculation of which polymorphic sites in the CHRM3 gene give the most significant contribution to the differences in phenotype. One regression model useful in the invention is described in PCT Application Serial No. PCT/US00/17540, entitled "Methods for Obtaining and Using Haplotype Data".

A second method for finding correlations between CHRM3 haplotype content and clinical responses uses predictive models based on error-minimizing optimization algorithms. One of many possible optimization algorithms is a genetic algorithm (R. Judson, "Genetic Algorithms and Their Uses in Chemistry" in Reviews in Computational Chemistry, Vol. 10, pp. 1-73, K. B. Lipkowitz and D. B. Boyd, eds. (VCH Publishers, New York, 1997). Simulated annealing (Press et al., "Numerical Recipes in C: The Art of Scientific Computing", Cambridge University Press (Cambridge) 1992, Ch. 10), neural networks (E. Rich and K. Knight, "Artificial Intelligence", 2nd Edition (McGraw-Hill, New York, 1991, Ch. 18), standard gradient descent methods (Press et al., supra Ch. 10), or other global or local optimization approaches (see discussion in Judson, supra) could also be used. Preferably, the correlation is found using a genetic algorithm approach as described in PCT Application Serial No. PCT/US00/17540.

Correlations may also be analyzed using analysis of variation (ANOVA) techniques to determine how much of the variation in the clinical data is explained by different subsets of the polymorphic sites in the CHRM3 gene. As described in PCT Application Serial No. PCT/US00/17540, ANOVA is used to test hypotheses about whether a response variable is caused by or correlated with one or more traits or variables that can be measured (Fisher and vanBelle, supra, Ch. 10).

From the analyses described above, a mathematical model may be readily constructed by the skilled artisan that predicts clinical response as a function of CHRM3 genotype or haplotype content. Preferably, the model is validated in one or more follow-up clinical trials designed to test the model.

The identification of an association between a clinical response and a genotype or haplotype (or haplotype pair) for the CHRM3 gene may be the basis for designing a diagnostic method to determine those individuals who will or will not respond to the treatment, or alternatively, will respond at a lower level and thus may require more treatment, i.e., a greater dose of a drug. The diagnostic method may take one of several forms: for example, a direct DNA test (i.e., genotyping or haplotyping one or more of the polymorphic sites in the CHRM3 gene), a serological test, or a physical exam measurement. The only requirement is that there be a good correlation between the diagnostic test results and the underlying CHRM3 genotype or haplotype that is in turn correlated with the clinical response. In a preferred embodiment, this diagnostic method uses the predictive haplotyping method described above.

Any or all analytical and mathematical operations involved in practicing the methods of the present invention may be implemented by a computer. In addition, the computer may execute a program that generates views (or screens) displayed on a display device and with which the user can interact to view and analyze large amounts of information relating to the CHRM3 gene and its genomic variation, including chromosome location, gene structure, and gene family, gene expression data, polymorphism data, genetic sequence data, and clinical data population data (e.g., data on ethnogeographic origin, clinical responses, genotypes, and haplotypes for one or more populations). The CHRM3 polymorphism data described herein may be stored as part of a relational database (e.g., an instance of an Oracle database or a set of ASCII flat files). These polymorphism data may be stored on the computer's hard drive or may, for example, be stored on a CD ROM or on one or more other storage devices accessible by the computer. For example, the data may be stored on one or more databases in communication with the computer via a network.

Preferred embodiments of the invention are described in the following examples. Other embodiments within the scope of the claims herein will be apparent to one skilled in the art from consideration of the specification or practice of the invention as disclosed herein. It is intended that the specification, together with the examples, be considered exemplary only, with the scope and spirit of the invention being indicated by the claims which follow the examples.

EXAMPLES

The Examples herein are meant to exemplify the various aspects of carrying out the invention and are not intended to limit the scope of the invention in any way. The Examples do not include detailed descriptions for conventional methods employed, such as in the performance of genomic DNA isolation, PCR and sequencing procedures. Such methods are well-known to those skilled in the art and are described in numerous publications, for example, Sambrook, Fritsch, and Maniatis, "Molecular

Cloning: A Laboratory Manual", 2nd Edition, Cold Spring Harbor Laboratory Press, USA, (1989).

EXAMPLE 1

This example illustrates examination of various regions of the CHRM3 gene for polymorphic sites.

Amplification of Target Regions

The following target regions of the CHRM3 gene were amplified using the PCR primer pairs listed below, with the sequences presented in the 5′ to 3′ direction and nucleotide positions shown for each region corresponding to the indicated GenBank Accession No.

Accession Number: Y00508.1	
Fragment 1	
Forward primer 32-85	GCTTCATAGAGATTCAGCACCCTG (SEQ ID NO:36)
Reverse primer Complement of 474-453	CACAATTACCAGGATGTTGCCG (SEQ ID NO:37)
PCR Product 413 nt	
Fragment 2	
Forward primer 300-321	CGTCACTCATTTCGGCAGCTAC (SEQ ID NO:38)
Reverse primer Complement of 854-832	GGAGGCACAGTTCTCTTTCCAAC (SEQ ID NO:39)
PCR Product 555 nt	
Fragment 3	
Forward primer 736-757	AAACGAACAACAAAGAGAGCCG (SEQ ID NO:40)
Reverse primer Complement of 1195-1174	TGTGGTCTTGGTCCATCTGCTC (SEQ ID NO:41)
PCR Product 460 nt	
Fragment 4	
Forward primer 1097-1116	GCATGAAACGCTCCAACAGG (SEQ ID NO:42)
Reverse primer Complement of 1635-1611	CCGCTTAGTGATCTGACTTCTGGTC (SEQ ID NO:43)
PCR Product 539 nt	
Fragment 5	
Forward primer 1505-1528	TGGACACAGCTAAGACTTCTGACG (SEQ ID NO:44)
Reverse primer Complement of 2011-1988	GTGTGCGTTTTGTCACTGCTATTG (SEQ ID NO:45)
PCR Product 507 nt	

These primer pairs were used in PCR reactions containing genomic DNA isolated from immortalized cell lines for each member of the Index Repository. The PCR reactions were carried out

under the following conditions:

Reaction volume	= 20 µl
10 x Advantage 2 Polymerase reaction buffer (Clontech)	= 2 µl
100 ng of human genomic DNA	= 1 µl
10 mM dNTP	= 0.4 µl
Advantage 2 Polymerase enzyme mix (Clontech)	= 0.2 µl
Forward Primer (10 µM)	= 0.4 µl
Reverse Primer (10 µM)	= 0.4 µl
Water	=15.6µl

Amplification profile:
94°C - 2 min. 1 cycle

94°C - 30 sec.
70°C - 45 sec. } 10 cycles
72°C - 1 min.

94°C - 30 sec.
64°C - 45 sec. } 35 cycles
72°C - 1 min.

Sequencing of PCR Products

The PCR products were purified by Solid Phase Reversible Immobilization using the protocol developed by the Whitehead Genome Center. A detailed protocol can be found at http://www.genome.wi.mit.edu/sequencing/protocols/pure/SPRI_pcr.html.

Briefly, five µl of carboxyl coated magnetic beads (10 mg/ml) and 60 µl of HYB BUFFER (2.5M NaCl/20% PEG 8000) were added to each PCR reaction mixture (20 µl). The reaction mixture was mixed well and incubated at room temperature (RT) for 10 min. The microtitre plate was placed on a magnet for 2 min and the beads washed twice with 150 µl of 70% EtOH. The beads were air dried for 2 min and the DNA was eluted in 25 µl of distilled water and incubated at RT for 5 min. The beads were magnetically separated and the supernatant removed for testing and sequencing.

The purified PCR products were sequenced in both directions using the primer sets described previously.

Analysis of Sequences for Polymorphic Sites

Sequences were analyzed for the presence of polymorphisms using the Polyphred program (Nickerson et al., Nucleic Acids Res. 14:2745-2751, 1997). The presence of a polymorphism was confirmed on both strands. The polymorphisms and their locations in the CHRM3 gene are listed in Table 3 below.

Table 3. Polymorphic Sites Identified in the CHRM3 Gene			
Polymorphic Site Number	Nucleotide Position	Reference Allele	Variant Allele
PS1	53(Acc#U29589.1)	T	C
PS2	394(Acc#U29589.1)	G	A
PS3	1387(Acc#U29589.1)	C	A
PS4	1493(Acc#U29589.1)	T	C

EXAMPLE 2

This example illustrates analysis of the CHRM3 polymorphisms identified in the Index Repository for human genotypes and haplotypes.

The different genotypes containing these polymorphisms that were observed in the reference population are shown in Table 4 below, with the haplotype pair indicating the combination of haplotypes determined for the individual using the haplotype derivation protocol described below. In Table 4, homozygous positions are indicated by one nucleotide and heterozygous positions are indicated by two nucleotides. Missing nucleotides in any given genotype in Table 4 were inferred based on linkage disequilibrium and/or Mendelian inheritance.

Table 4. Genotypes and Haplotype Pairs Observed for the CHRM3 Gene						
Genotype Number	Polymorphic Sites				HAP Pair	
	PS1	PS2	PS3	PS4		
1	T	G	C	T	3	3
2	C	G	C	T	1	1
3	T	G	C/A	T/C	3	2
4	T/C	G	C	T	3	1

The haplotype pairs shown in Table 4 were estimated from the unphased genotypes using an extension of Clark's algorithm (Clark, A.G. (1990) *Mol Bio Evol* 7, 111-122), as described in U.S. Provisional Application Serial No. 60/198,340 entitled "A Method and System for Determining Haplotypes from a Collection of Polymorphisms". In this method, haplotypes are assigned directly from individuals who are homozygous at all sites or heterozygous at no more than one of the variable sites. In our analysis, the list of haplotypes is augmented with haplotypes obtained from a three-generation Caucasian family and a two-generation African-American family. This list of haplotypes is then used to deconvolute the unphased genotypes in the remaining (multiply heterozygous) individuals.

By following this protocol, it was determined that the Index Repository examined herein and, by extension, the general population contains the 3 human CHRM3 haplotypes shown in Table 5 below.

Table 5. Haplotypes Identified in the CHRM3 Gene				
	Polymorphic Sites			
Haplotype Number	PS 1	PS 2	PS 3	PS 4
1	C	G	C	T
2	T	G	A	C
3	T	G	C	T

In view of the above, it will be seen that the several advantages of the invention are achieved and other advantageous results attained.

As various changes could be made in the above methods and compositions without departing from the scope of the invention, it is intended that all matter contained in the above description and shown in the accompanying drawings shall be interpreted as illustrative and not in a limiting sense.

All references cited in this specification, including patents and patent applications, are hereby incorporated in their entirety by reference. The discussion of references herein is intended merely to summarize the assertions made by their authors and no admission is made that any reference constitutes prior art. Applicants reserve the right to challenge the accuracy and pertinency of the cited references.

What is Claimed is:

1. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of:
   (a) a first nucleotide sequence which is a polymorphic variant of a reference sequence for cholinergic receptor, muscarinic 3 (CHRM3) gene or a fragment thereof, wherein the reference sequence comprises SEQ ID NO:1and the polymorphic variant comprises at least one polymorphism selected from the group consisting of cytosine at PS1, adenine at PS2, adenine at PS3 and cytosine at PS4; and
   (b) a second nucleotide sequence which is complementary to the first nucleotide sequence.
2. The isolated polynucleotide of claim 1 which comprises a CHRM3 isogene.
3. The isolated polynucleotide of claim 1 which is a DNA molecule and comprises both the first and second nucleotide sequences and further comprises expression regulatory elements operably linked to the first nucleotide sequence.
4. A recombinant nonhuman organism transformed or transfected with the isolated polynucleotide of claim 1, wherein the organism expresses a CHRM3 protein encoded by the first nucleotide sequence.
5. The recombinant organism of claim 4 which is a nonhuman transgenic animal.
6. The isolated polynucleotide of claim 1, wherein the first nucleotide sequence is a polymorphic variant of a fragment of the CHRM3 gene, the fragment comprising one or more polymorphisms selected from the group consisting of cytosine at PS1, adenine at PS2, adenine at PS3 and cytosine at PS4.
7. An isolated polynucleotide comprising a nucleotide sequence which is a polymorphic variant of a reference sequence for the CHRM3 cDNA or a fragment thereof, wherein the reference sequence comprises SEQ ID NO:2 and the polymorphic variant comprises at least one polymorphism selected from the group consisting of adenine at a position corresponding to nucleotide 193, adenine at a position corresponding to nucleotide 1186 and cytosine at a position corresponding to nucleotide 1292.
8. A recombinant nonhuman organism transformed or transfected with the isolated polynucleotide of claim 7, wherein the organism expresses a cholinergic receptor, muscarinic 3 (CHRM3) protein encoded by the polymorphic variant sequence.
9. The recombinant organism of claim 8 which is a nonhuman transgenic animal.
10. An isolated polypeptide comprising an amino acid sequence which is a polymorphic variant of a reference sequence for the CHRM3 protein or a fragment thereof, wherein the reference sequence comprises SEQ ID NO: 3 and the polymorphic variant comprises one or more variant amino acids selected from the group consisting of isoleucine at a position corresponding to amino acid position 65, methionine at a position corresponding to amino acid position 396, and proline at a position corresponding to amino acid position 431.
11. An isolated antibody specific for and immunoreactive with the isolated polypeptide of claim 10.

12. A method for screening for drugs targeting the isolated polypeptide of claim 10 which comprises contacting the CHRM3 polymorphic variant with a candidate agent and assaying for binding activity.

13. A composition comprising at least one genotyping oligonucleotide for detecting a polymorphism in the cholinergic receptor, muscarinic 3 (CHRM3) gene at a polymorphic site selected from PS1, PS2, PS3, and PS4.

14. The composition of claim 13, wherein the genotyping oligonucleotide is an allele-specific oligonucleotide that specifically hybridizes to an allele of the CHRM3 gene at a region containing the polymorphic site.

15. The composition of claim 14, wherein the allele-specific oligonucleotide comprises a nucleotide sequence selected from the group consisting of of SEQ ID NOS:4-11, the complements of SEQ ID NOS: 4-11, and SEQ ID NOS:12-27.

16. The composition of claim 13, wherein the genotyping oligonucleotide is a primer-extension oligonucleotide.

17. A method for genotyping the cholinergic receptor, muscarinic 3 (CHRM3) gene of an individual, comprising determining for the two copies of the CHRM3 gene present in the individual the identity of the nucleotide pair at one or more polymorphic sites (PS) selected from PS1, PS2, PS3, and PS4.

18. The method of claim 17, wherein the determining step comprises:

(a) isolating from the individual a nucleic acid mixture comprising both copies of the CHRM3 gene, or a fragment thereof, that are present in the individual;

(b) amplifying from the nucleic acid mixture a target region containing at least one of the polymorphic sites;

(c) hybridizing a primer extension oligonucleotide to one allele of the amplified target region;

(d) performing a nucleic acid template-dependent, primer extension reaction on the hybridized genotyping oligonucleotide in the presence of at least two different terminators of the reaction, wherein said terminators are complementary to the alternative nucleotides present at the polymorphic site; and

(e) detecting the presence and identity of the terminator in the extended genotyping oligonucleotide.

19. A method for haplotyping the cholinergic receptor, muscarinic 3 (CHRM3) gene of an individual which comprises determining, for one copy of the CHRM3 gene present in the individual, the identity of the nucleotide at one or more polymorphic sites (PS) selected from PS1, PS2, PS3, and PS4.

20. The method of claim 19, wherein the determining step comprises

(a) isolating from the individual a nucleic acid molecule containing only one of the two copies of

the CHRM3 gene, or a fragment thereof, that is present in the individual;

(b) amplifying from the nucleic acid molecule a target region containing at least one of the polymorphic sites;

(c) hybridizing a primer extension oligonucleotide to one allele of the amplified target region;

(d) performing a nucleic acid template-dependent, primer extension reaction on the hybridized genotyping oligonucleotide in the presence of at least two different terminators of the reaction, wherein said terminators are complementary to the alternative nucleotides present at the polymorphic site; and

(e) detecting the presence and identity of the terminator in the extended genotyping oligonucleotide.

21. A method for predicting a haplotype pair for the cholinergic receptor, muscarinic 3 (CHRM3) gene of an individual comprising:

(a) identifying an CHRM3 genotype for the individual at two or more of polymorphic sites selected from PS1, PS2, PS3, and PS4;

(b) enumerating all possible haplotype pairs which are consistent with the genotype;

(c) accessing data containing the CHRM3 haplotype pairs determined in a reference population; and

(d) assigning a haplotype pair to the individual that is consistent with the data.

22. A method for identifying an association between a trait and at least one genotype or haplotype of the cholinergic receptor, muscarinic 3 gene which comprises comparing the frequency of the genotype or haplotype in a population exhibiting the trait with the frequency of the genotype or haplotype in a reference population, wherein the genotype or haplotype comprises a nucleotide pair or nucleotide located at one or more polymorphic sites selected from PS1, PS2, PS3, and PS4, wherein a higher frequency of the genotype or haplotype in the trait population than in the reference population indicates the trait is associated with the genotype or haplotype.

23 The method of claim 22, wherein the haplotype is selected from haplotype numbers 1-3 shown in Table 5.

24. The method of claim 23, wherein the trait is a clinical response to a drug targeting CHRM3.

25. A computer system for storing and analyzing polymorphism data for the cholinergic receptor, muscarinic 3 gene, comprising:

(a) a central processing unit (CPU);

(b) a communication interface;

(c) a display device;

(d) an input device; and

(e) a database containing the polymorphism data;

wherein the polymorphism data comprises the genotypes and haplotype pairs shown in Table 4

and the haplotypes shown in Table 5.

26. A genome anthology for the cholinergic receptor, muscarinic 3 (CHRM3) gene which comprises CHRM3 isogenes defined by haplotypes 1-3 shown in Table 5.

POLYMORPHISMS IN THE CHRM3 GENE

```
ATCGATGTCT GTCTGCCCTA GACTCCACTT ATTTAAAATA AGAGAATGAA
CTTGATGTTT GGCTTCATAG AGATTCAGCA CCCTGTAATA GGCCTTCCAT      100
  C
GTCTTTTAAC GTATGTAATG CAAAGAACAA ACAAATAAAG GCAGAAATTT
TTCTAACTCT GTCTCTTCTC TCTTTCCCCC AGACTATGTC AGAGAGTCAC      200
AATGACCTTG CACAATAACA GTACAACCTC GCCTTTGTTT CCAAACATCA
 [exon 1: 201..
GCTCCTCCTG GATACACAGC CCCTCCGATG CAGGGCTGCC CCCGGGAACC      300
GTCACTCATT TCGGCAGCTA CAATGTTTCT CGAGCAGCTG GCAATTTCTC
CTCTCCAGAC GGTACCACCG ATGACCCTCT GGGAGGTCAT ACCGTCTGGC      400
                                                  A
AAGTGGTCTT CATCGCTTTC TTAACGGGCA TCCTGGCCTT GGTGACCATC
ATCGGCAACA TCCTGGTAAT TGTGTCATTT AAGGTCAACA AGCAGCTGAA      500
GACGGTCAAC AACTACTTCC TCTTAAGCCT GGCCTGTGCC GATCTGATTA
TCGGGGTCAT TTCAATGAAT CTGTTTACGA CCTACATCAT CATGAATCGA      600
TGGGCCTTAG GGAACTTGGC CTGTGACCTC TGGCTTGCCA TTGACTACGT
AGCCAGCAAT GCCTCTGTTA TGAATCTTCT GGTCATCAGC TTTGACAGAT      700
ACTTTTCCAT CACGAGGCCG CTCACGTACC GAGCCAAACG AACAACAAAG
AGAGCCGGTG TGATGATCGG TCTGGCTTGG GTCATCTCCT TTGTCCTTTG      800
GGCTCCTGCC ATCTTGTTCT GGCAATACTT TGTTGGAAAG AGAACTGTGC
CTCCGGGAGA GTGCTTCATT CAGTTCCTCA GTGAGCCCAC CATTACTTTT      900
GGCACAGCCA TCGCTGCTTT TTATATGCCT GTCACCATTA TGACTATTTT
ATACTGGAGG ATCTATAAGG AAACTGAAAA GCGTACCAAA GAGCTTGCTG     1000
GCCTGCAAGC CTCTGGGACA GAGGCAGAGA CAGAAAACTT TGTCCACCCC
ACGGGCAGTT CTCGAAGCTG CAGCAGTTAC GAACTTCAAC AGCAAAGCAT     1100
GAAACGCTCC AACAGGAGGA AGTATGGCCG CTGCCACTTC TGGTTCACAA
CCAAGAGCTG GAAACCCAGC TCCGAGCAGA TGGACCAAGA CCACAGCAGC     1200
AGTGACAGTT GGAACAACAA TGATGCTGCT GCCTCCCTGG AGAACTCCGC
CTCCTCCGAC GAGGAGGACA TTGGCTCCGA GACGAGAGCC ATCTACTCCA     1300
TCGTGCTCAA GCTTCCGGGT CACAGCACCA TCCTCAACTC CACCAAGTTA
CCCTCATCGG ACAACCTGCA GGTGCCTGAG GAGGAGCTGG GGATGGTGGA     1400
                                     A
CTTGGAGAGG AAAGCCGACA AGCTGCAGGC CCAGAAGAGC GTGGACGATG
GAGGCAGTTT TCCAAAAAGC TTCTCCAAGC TTCCCATCCA GCTAGAGTCA     1500
                                                  C
GCCGTGGACA CAGCTAAGAC TTCTGACGTC AACTCCTCAG TGGGTAAGAG
CACGGCCACT CTACCTCTGT CCTTCAAGGA AGCCACTCTG GCCAAGAGGT     1600
TTGCTCTGAA GACCAGAAGT CAGATCACTA AGCGGAAAAG GATGTCCCTG
GTCAAGGAGA AGAAAGCGGC CCAGACCCTC AGTGCGATCT TGCTTGCCTT     1700
CATCATCACT TGGACCCCAT ACAACATCAT GGTTCTGGTG AACACCTTTT
GTGACAGCTG CATACCCAAA ACCTTTTGGA ATCTGGGCTA CTGGCTGTGC     1800
TACATCAACA GCACCGTGAA CCCCGTGTGC TATGCTCTGT GCAACAAAAC
ATTCAGAACC ACTTTCAAGA TGCTGCTGCT GTGCCAGTGT GACAAAAAAA     1900
AGAGGCGCAA GCAGCAGTAC CAGCAGAGAC AGTCGGTCAT TTTTCACAAG
CGCGCACCCG AGCAGGCCTT GTAGAATGAG GTTGTATCAA TAGCAGTGAC     2000
                ..1974]
AAAACGCACA CATCAACCCA CAGACCTTAG GAGGAGGAAG GCGAGGGCGG
GGTGACTTCT GGTGATGATA AAAATGGTTT TATCACCCAG ATGTGAAAGA     2100
AGCTGCCTGT TTACTGATCC ATTGAATAAA CCCATTTTAA TAGAAAAAGT
CAATACCAAT TCAGCAAAAA GAAAAAAAAA ACATACTACT GAATATAAAG     2200
```

FIGURE 1A

```
AAATTTATTC TGAAATAGAC TTTACGTGTT TTTTTCTTAA AGAGGAGAAA
AATATTGCTT GACGGCAATT ATATACCCAA AGTGATTTGC CTGGGTCCTT      2300
TAATTCCCAT TAGCTTTGGA ATCTCAGATG AGCATAGCTG ACCCAGTTCC
CACATTCTTC CCAAGGATCC AAAAGTGGGA ATCCAGACCC CAAGTGGAAC      2400
ACTGCAGGCT TACGAATCTG TGGTTCCAAA ATTATTTCAT ACGTTGCAAA
GCTGAATCTT CTTGTCCCAA TAGAGCTTCC TGTCTTTTCT TTGGTGTGTT      2500
GTTAAACTCT ATTTGTGGAC TTGATTCTTG ATTCTTGCAA AGTACTGTTT
TGTGCAGTTC AAGTTTCGTA CAAATAAAAT ACTTAAGTAT ATATATATGT      2600
GTGAGTTCTG CACGCACACA CATAGTGTAT ATAATATCAT GGGAAACACT
GAACTGGCAA ATTATTCCTG CAACATACGC TTTCAGTACT TTGGTAACTG      2700
AAGTTCTCTA GGATCCTAAT GCAACATTAA CGTGAAATAA GCCCAGTGTA
ATGTTTTTGC AAACCAGGGC TGTTTTCCAC AGAGAGCAGC CAGGCCTTCC      2800
CAGCAGGTCT GTGCAGAGCG GACAGGCTCG TGAGTCAGCT GAGCGCCGTG
GCTTCGCCAG ACTTGGTGTT AAGCAACCTC CTTTGTTGAT GTCTCAACAG      2900
AGCTAAATCG GGGCCCCTCT GAGCTCAAAG AATGAACCAC ATCCACACGT
TTGAATTTAA TCATCTAAAT CTGAATGTTT CAGAACAAAA TTTCTGCTAT      3000
CTAAACTGCT TGAAACTCAA TAATAGTGTC ACGTTTGAAT GTCATACACA
GCAATATATA TATATGTGTA TATATATATA TATGGCAAAG CAAAAAAAAA      3100
AACATGGTAA GAGAGAATGA AGGAGAACAT TGTGTTTGAT TCTTGCTGAA
TGGCACCTTC TCAAAGAAAA TAGGGCTTGC ACCTTTGTTA ATCAGCTGTG      3200
GCCAGTGCTT TCTGGTGTTC ATTGTGTAAC CTTCACCCAG GAATAGGTGA
GGTTTTAGGA AGTTACATGT CCTCTGAAGA AAGAATTACA CTCTGAAAAG      3300
TAATGCTTCA AATTGATTTC CTTACCTTTT GGGAAAAAAA AAAAATTGTT
TTTTTGCATT CTCCCTTGAA TTGACCAAAA TGTTAACTGT TTCATTTGGG      3400
GAGGGGATGG GGTGCTGCCA TCATTGTCGT TGTTGTTGCT GCTGTAGCTG
TTGGGGTTTC TTTTCCTGTT GCCGGGGCTG TTTGGGGAGA GGGAGGGGAG      3500
GGAGGTGGGA GGGCCGCGGA GATATCTTCC C                           3531
```

FIGURE 1B

POLYMORPHISMS IN THE CODING SEQUENCE OF CHRM3

```
ATGACCTTGC ACAATAACAG TACAACCTCG CCTTTGTTTC CAAACATCAG
CTCCTCCTGG ATACACAGCC CCTCCGATGC AGGGCTGCCC CCGGGAACCG      100
TCACTCATTT CGGCAGCTAC AATGTTTCTC GAGCAGCTGG CAATTTCTCC
TCTCCAGACG GTACCACCGA TGACCCTCTG GGAGGTCATA CCGTCTGGCA      200
                                                  A
AGTGGTCTTC ATCGCTTTCT TAACGGGCAT CCTGGCCTTG GTGACCATCA
TCGGCAACAT CCTGGTAATT GTGTCATTTA AGGTCAACAA GCAGCTGAAG      300
ACGGTCAACA ACTACTTCCT CTTAAGCCTG GCCTGTGCCG ATCTGATTAT
CGGGGTCATT TCAATGAATC TGTTTACGAC CTACATCATC ATGAATCGAT      400
GGGCCTTAGG GAACTTGGCC TGTGACCTCT GGCTTGCCAT TGACTACGTA
GCCAGCAATG CCTCTGTTAT GAATCTTCTG GTCATCAGCT TTGACAGATA      500
CTTTTCCATC ACGAGGCCGC TCACGTACCG AGCCAAACGA ACAACAAAGA
GAGCCGGTGT GATGATCGGT CTGGCTTGGG TCATCTCCTT TGTCCTTTGG      600
GCTCCTGCCA TCTTGTTCTG GCAATACTTT GTTGGAAAGA GAACTGTGCC
TCCGGGAGAG TGCTTCATTC AGTTCCTCAG TGAGCCCACC ATTACTTTTG      700
GCACAGCCAT CGCTGCTTTT TATATGCCTG TCACCATTAT GACTATTTTA
TACTGGAGGA TCTATAAGGA AACTGAAAAG CGTACCAAAG AGCTTGCTGG      800
CCTGCAAGCC TCTGGGACAG AGGCAGAGAC AGAAAACTTT GTCCACCCCA
CGGGCAGTTC TCGAAGCTGC AGCAGTTACG AACTTCAACA GCAAAGCATG      900
AAACGCTCCA ACAGGAGGAA GTATGGCCGC TGCCACTTCT GGTTCACAAC
CAAGAGCTGG AAACCCAGCT CCGAGCAGAT GGACCAAGAC CACAGCAGCA     1000
GTGACAGTTG GAACAACAAT GATGCTGCTG CCTCCCTGGA GAACTCCGCC
TCCTCCGACG AGGAGGACAT TGGCTCCGAG ACGAGAGCCA TCTACTCCAT     1100
CGTGCTCAAG CTTCCGGGTC ACAGCACCAT CCTCAACTCC ACCAAGTTAC
CCTCATCGGA CAACCTGCAG GTGCCTGAGG AGGAGCTGGG GATGGTGGAC     1200
                                     A
TTGGAGAGGA AAGCCGACAA GCTGCAGGCC CAGAAGAGCG TGGACGATGG
AGGCAGTTTT CCAAAAAGCT TCTCCAAGCT TCCCATCCAG CTAGAGTCAG     1300
                                                  C
CCGTGGACAC AGCTAAGACT TCTGACGTCA ACTCCTCAGT GGGTAAGAGC
ACGGCCACTC TACCTCTGTC CTTCAAGGAA GCCACTCTGG CCAAGAGGTT     1400
TGCTCTGAAG ACCAGAAGTC AGATCACTAA GCGGAAAAGG ATGTCCCTGG
TCAAGGAGAA GAAAGCGGCC CAGACCCTCA GTGCGATCTT GCTTGCCTTC     1500
ATCATCACTT GGACCCCATA CAACATCATG GTTCTGGTGA ACACCTTTTG
TGACAGCTGC ATACCCAAAA CCTTTTGGAA TCTGGGCTAC TGGCTGTGCT     1600
ACATCAACAG CACCGTGAAC CCCGTGTGCT ATGCTCTGTG CAACAAAACA
TTCAGAACCA CTTTCAAGAT GCTGCTGCTG TGCCAGTGTG ACAAAAAAAA     1700
GAGGCGCAAG CAGCAGTACC AGCAGAGACA GTCGGTCATT TTTCACAAGC
GCGCACCCGA GCAGGCCTTG TAG                                  1773
```

FIGURE 2

ISOFORMS OF THE CHRM3 PROTEIN

```
MTLHNNSTTS PLFPNISSSW IHSPSDAGLP PGTVTHFGSY NVSRAAGNFS
SPDGTTDDPL GGHTVWQVVF IAFLTGILAL VTIIGNILVI VSFKVNKQLK      100
               I
TVNNYFLLSL ACADLIIGVI SMNLFTTYII MNRWALGNLA CDLWLAIDYV
ASNASVMNLL VISFDRYFSI TRPLTYRAKR TTKRAGVMIG LAWVISFVLW      200
APAILFWQYF VGKRTVPPGE CFIQFLSEPT ITFGTAIAAF YMPVTIMTIL
YWRIYKETEK RTKELAGLQA SGTEAETENF VHPTGSSRSC SSYELQQQSM      300
KRSNRRKYGR CHFWFTTKSW KPSSEQMDQD HSSSDSWNNN DAAASLENSA
SSDEEDIGSE TRAIYSIVLK LPGHSTILNS TKLPSSDNLQ VPEEELGMVD      400
                                                    M
LERKADKLQA QKSVDDGGSF PKSFSKLPIQ LESAVDTAKT SDVNSSVGKS
                                 P
TATLPLSFKE ATLAKRFALK TRSQITKRKR MSLVKEKKAA QTLSAILLAF      500
IITWTPYNIM VLVNTFCDSC IPKTFWNLGY WLCYINSTVN PVCYALCNKT
FRTTFKMLLL CQCDKKKRRK QQYQQRQSVI FHKRAPEQAL                 590
```

FIGURE 3

SEQUENCE LISTING

<110> Genaissance Pharmaceuticals
Nandabalan, Krishnan
Denton, R. Rex
Stephens, J. Claiborne
Choi, Julie Y

<120> DRUG TARGET ISOGENES: POLYMORPHISMS IN THE CHOLINERGIC RECEPTOR, MUSCARINIC 3 GENE

<130> MWH-0018 PCT CHRM3

<140> TBA
<141> 2000-10-12

<150> 60/159,860
<151> 1999-10-15

<160> 45

<170> PatentIn Ver. 2.1

<210> 1
<211> 3531
<212> DNA
<213> Homo sapiens

<400> 1
atcgatgtct gtctgcccta gactccactt atttaaaata agagaatgaa cttgatgttt 60
ggcttcatag agattcagca ccctgtaata ggccttccat gtcttttaac gtatgtaatg 120
caaagaacaa acaaataaag gcagaaattt ttctaactct gtctcttctc tctttccccc 180
agactatgtc agagagtcac aatgaccttg cacaataaca gtacaacctc gcctttgttt 240
ccaaacatca gctcctcctg gatacacagc ccctccgatg cagggctgcc cccgggaacc 300
gtcactcatt tcggcagcta caatgtttct cgagcagctg gcaatttctc ctctccagac 360
ggtaccaccg atgaccctct gggaggtcat accgtctggc aagtggtctt catcgctttc 420
ttaacgggca tcctggcctt ggtgaccatc atcggcaaca tcctggtaat tgtgtcattt 480
aaggtcaaca agcagctgaa gacggtcaac aactacttcc tcttaagcct ggcctgtgcc 540
gatctgatta tcggggtcat ttcaatgaat ctgtttacga cctacatcat catgaatcga 600
tgggccttag ggaacttggc ctgtgacctc tggcttgcca ttgactacgt agccagcaat 660
gcctctgtta tgaatcttct ggtcatcagc tttgacagat acttttccat cacgaggccg 720
ctcacgtacc gagccaaacg aacaacaaag agagccggtg tgatgatcgg tctggcttgg 780
gtcatctcct ttgtcctttg ggctcctgcc atcttgttct ggcaatactt tgttggaaag 840
agaactgtgc ctccgggaga gtgcttcatt cagttcctca gtgagcccac cattactttt 900
ggcacagcca tcgctgcttt ttatatgcct gtcaccatta tgactatttt atactggagg 960
atctataagg aaactgaaaa gcgtaccaaa gagcttgctg gcctgcaagc ctctgggaca 1020
gaggcagaga cagaaaactt tgtccacccc acgggcagtt ctcgaagctg cagcagttac 1080
gaacttcaac agcaaagcat gaaacgctcc aacaggagga agtatggccg ctgccacttc 1140

tggttcacaa ccaagagctg gaaacccagc tccgagcaga tggaccaaga ccacagcagc 1200
agtgacagtt ggaacaacaa tgatgctgct gcctccctgg agaactccgc ctcctccgac 1260
gaggaggaca ttggctccga gacgagagcc atctactcca tcgtgctcaa gcttccgggt 1320
cacagcacca tcctcaactc caccaagtta ccctcatcgg acaacctgca ggtgcctgag 1380
gaggagctgg ggatggtgga cttggagagg aaagccgaca agctgcaggc ccagaagagc 1440
gtggacgatg gaggcagttt tccaaaaagc ttctccaagc ttcccatcca gctagagtca 1500
gccgtggaca cagctaagac ttctgacgtc aactcctcag tgggtaagag cacggccact 1560
ctacctctgt ccttcaagga agccactctg gccaagaggt ttgctctgaa gaccagaagt 1620
cagatcacta agcggaaaag gatgtccctg gtcaaggaga agaaagcggc ccagaccctc 1680
agtgcgatct tgcttgcctt catcatcact tggaccccat acaacatcat ggttctggtg 1740
aacacctttt gtgacagctg catacccaaa acctttgga atctgggcta ctggctgtgc 1800
tacatcaaca gcaccgtgaa ccccgtgtgc tatgctctgt gcaacaaaac attcagaacc 1860
actttcaaga tgctgctgct gtgccagtgt gacaaaaaaa agaggcgcaa gcagcagtac 1920
cagcagagac agtcggtcat ttttcacaag cgcgcacccg agcaggcctt gtagaatgag 1980
gttgtatcaa tagcagtgac aaaacgcaca catcaaccca cagaccttag gaggaggaag 2040
gcgagggcgg ggtgacttct ggtgatgata aaaatggttt tatcacccag atgtgaaaga 2100
agctgcctgt ttactgatcc attgaataaa cccattttaa tagaaaaagt caataccaat 2160
tcagcaaaaa gaaaaaaaaa acatactact gaatataaag aaatttattc tgaaatagac 2220
tttacgtgtt tttttcttaa agaggagaaa aatattgctt gacggcaatt atatacccaa 2280
agtgatttgc ctgggtcctt taattcccat tagctttgga atctcagatg agcatagctg 2340
acccagttcc cacattcttc ccaaggatcc aaaagtggga atccagaccc caagtggaac 2400
actgcaggct tacgaatctg tggttccaaa attatttcat acgttgcaaa gctgaatctt 2460
cttgtcccaa tagagcttcc tgtcttttct ttggtgtgtt gttaaactct atttgtggac 2520
ttgattcttg attcttgcaa agtactgttt tgtgcagttc aagtttcgta caaataaaat 2580
acttaagtat atatatatgt gtgagttctg cacgcacaca catagtgtat ataatatcat 2640
gggaaacact gaactggcaa attattcctg caacatacgc tttcagtact ttggtaactg 2700
aagttctcta ggatcctaat gcaacattaa cgtgaaataa gcccagtgta atgtttttgc 2760
aaaccagggc tgttttccac agagagcagc caggccttcc cagcaggtct gtgcagagcg 2820
gacaggctcg tgagtcagct gagcgccgtg gcttcgccag acttggtgtt aagcaacctc 2880
ctttgttgat gtctcaacag agctaaatcg gggcccctct gagctcaaag aatgaaccac 2940
atccacacgt ttgaatttaa tcatctaaat ctgaatgttt cagaacaaaa tttctgctat 3000
ctaaactgct tgaaactcaa taatagtgtc acgtttgaat gtcatacaca gcaatatata 3060
tatatgtgta tatatatata tatggcaaag caaaaaaaaa aacatggtaa gagagaatga 3120
aggagaacat tgtgtttgat tcttgctgaa tggcaccttc tcaaagaaaa tagggcttgc 3180
acctttgtta atcagctgtg gccagtgctt tctggtgttc attgtgtaac cttcacccag 3240
gaataggtga ggttttagga agttacatgt cctctgaaga aagaattaca ctctgaaaag 3300
taatgcttca aattgatttc cttacctttt gggaaaaaaa aaaaattgtt tttttgcatt 3360
ctcccttgaa ttgaccaaaa tgttaactgt ttcatttggg gaggggatgg ggtgctgcca 3420
tcattgtcgt tgttgttgct gctgtagctg ttggggtttc ttttcctgtt gccggggctg 3480
tttggggaga gggaggggag ggaggtggga gggccgcgga gatatcttcc c 3531

<210> 2
<211> 1773
<212> DNA
<213> Homo sapiens

<400> 2

atgaccttgc acaataacag tacaacctcg cctttgtttc caaacatcag ctcctcctgg 60
atacacagcc cctccgatgc agggctgccc ccgggaaccg tcactcattt cggcagctac 120
aatgtttctc gagcagctgg caatttctcc tctccagacg gtaccaccga tgaccctctg 180
ggaggtcata ccgtctggca agtggtcttc atcgctttct taacgggcat cctggccttg 240
gtgaccatca tcggcaacat cctggtaatt gtgtcattta aggtcaacaa gcagctgaag 300
acggtcaaca actacttcct cttaagcctg gcctgtgccg atctgattat cggggtcatt 360
tcaatgaatc tgtttacgac ctacatcatc atgaatcgat gggccttagg gaacttggcc 420
tgtgacctct ggcttgccat tgactacgta gccagcaatg cctctgttat gaatcttctg 480
gtcatcagct ttgacagata cttttccatc acgaggccgc tcacgtaccg agccaaacga 540
acaacaaaga gagccggtgt gatgatcggt ctggcttggg tcatctcctt tgtcctttgg 600
gctcctgcca tcttgttctg gcaatacttt gttggaaaga gaactgtgcc tccgggagag 660
tgcttcattc agttcctcag tgagcccacc attacttttg gcacagccat cgctgctttt 720
tatatgcctg tcaccattat gactatttta tactggagga tctataagga aactgaaaag 780
cgtaccaaag agcttgctgg cctgcaagcc tctgggacag aggcagagac agaaaacttt 840
gtccacccca cgggcagttc tcgaagctgc agcagttacg aacttcaaca gcaaagcatg 900
aaacgctcca acaggaggaa gtatggccgc tgccacttct ggttcacaac caagagctgg 960
aaacccagct ccgagcagat ggaccaagac cacagcagca gtgacagttg gaacaacaat 1020
gatgctgctg cctccctgga gaactccgcc tcctccgacg aggaggacat tggctccgag 1080
acgagagcca tctactccat cgtgctcaag cttccgggtc acagcaccat cctcaactcc 1140
accaagttac cctcatcgga caacctgcag gtgcctgagg aggagctggg gatggtggac 1200
ttggagagga aagccgacaa gctgcaggcc cagaagagcg tggacgatgg aggcagtttt 1260
ccaaaaagct tctccaagct tcccatccag ctagagtcag ccgtggacac agctaagact 1320
tctgacgtca actcctcagt gggtaagagc acggccactc tacctctgtc cttcaaggaa 1380
gccactctgg ccaagaggtt tgctctgaag accagaagtc agatcactaa gcggaaaagg 1440
atgtccctgg tcaaggagaa gaaagcggcc cagaccctca gtgcgatctt gcttgccttc 1500
atcatcactt ggacccctata caacatcatg gttctggtga acacctttg tgacagctgc 1560
atacccaaaa ccttttggaa tctgggctac tggctgtgct acatcaacag caccgtgaac 1620
cccgtgtgct atgctctgtg caacaaaaca ttcagaacca ctttcaagat gctgctgctg 1680
tgccagtgtg acaaaaaaaa gaggcgcaag cagcagtacc agcagagaca gtcggtcatt 1740
tttcacaagc gcgcacccga gcaggccttg tag 1773

<210> 3
<211> 590
<212> PRT
<213> Homo sapiens

<400> 3
Met Thr Leu His Asn Asn Ser Thr Thr Ser Pro Leu Phe Pro Asn Ile
1 5 10 15

Ser Ser Ser Trp Ile His Ser Pro Ser Asp Ala Gly Leu Pro Pro Gly
20 25 30

Thr Val Thr His Phe Gly Ser Tyr Asn Val Ser Arg Ala Ala Gly Asn
35 40 45

Phe Ser Ser Pro Asp Gly Thr Thr Asp Asp Pro Leu Gly Gly His Thr

```
     50                  55                  60

Val Trp Gln Val Val Phe Ile Ala Phe Leu Thr Gly Ile Leu Ala Leu
 65                  70                  75                  80

Val Thr Ile Ile Gly Asn Ile Leu Val Ile Val Ser Phe Lys Val Asn
                85                  90                  95

Lys Gln Leu Lys Thr Val Asn Asn Tyr Phe Leu Leu Ser Leu Ala Cys
            100                 105                 110

Ala Asp Leu Ile Ile Gly Val Ile Ser Met Asn Leu Phe Thr Thr Tyr
        115                 120                 125

Ile Ile Met Asn Arg Trp Ala Leu Gly Asn Leu Ala Cys Asp Leu Trp
    130                 135                 140

Leu Ala Ile Asp Tyr Val Ala Ser Asn Ala Ser Val Met Asn Leu Leu
145                 150                 155                 160

Val Ile Ser Phe Asp Arg Tyr Phe Ser Ile Thr Arg Pro Leu Thr Tyr
                165                 170                 175

Arg Ala Lys Arg Thr Thr Lys Arg Ala Gly Val Met Ile Gly Leu Ala
            180                 185                 190

Trp Val Ile Ser Phe Val Leu Trp Ala Pro Ala Ile Leu Phe Trp Gln
        195                 200                 205

Tyr Phe Val Gly Lys Arg Thr Val Pro Pro Gly Glu Cys Phe Ile Gln
    210                 215                 220

Phe Leu Ser Glu Pro Thr Ile Thr Phe Gly Thr Ala Ile Ala Ala Phe
225                 230                 235                 240

Tyr Met Pro Val Thr Ile Met Thr Ile Leu Tyr Trp Arg Ile Tyr Lys
                245                 250                 255

Glu Thr Glu Lys Arg Thr Lys Glu Leu Ala Gly Leu Gln Ala Ser Gly
            260                 265                 270

Thr Glu Ala Glu Thr Glu Asn Phe Val His Pro Thr Gly Ser Ser Arg
        275                 280                 285

Ser Cys Ser Ser Tyr Glu Leu Gln Gln Gln Ser Met Lys Arg Ser Asn
    290                 295                 300

Arg Arg Lys Tyr Gly Arg Cys His Phe Trp Phe Thr Thr Lys Ser Trp
```

```
305                 310                 315                 320

Lys Pro Ser Ser Glu Gln Met Asp Gln Asp His Ser Ser Ser Asp Ser
                325                 330                 335

Trp Asn Asn Asn Asp Ala Ala Ala Ser Leu Glu Asn Ser Ala Ser Ser
            340                 345                 350

Asp Glu Glu Asp Ile Gly Ser Glu Thr Arg Ala Ile Tyr Ser Ile Val
        355                 360                 365

Leu Lys Leu Pro Gly His Ser Thr Ile Leu Asn Ser Thr Lys Leu Pro
    370                 375                 380

Ser Ser Asp Asn Leu Gln Val Pro Glu Glu Glu Leu Gly Met Val Asp
385                 390                 395                 400

Leu Glu Arg Lys Ala Asp Lys Leu Gln Ala Gln Lys Ser Val Asp Asp
                405                 410                 415

Gly Gly Ser Phe Pro Lys Ser Phe Ser Lys Leu Pro Ile Gln Leu Glu
            420                 425                 430

Ser Ala Val Asp Thr Ala Lys Thr Ser Asp Val Asn Ser Ser Val Gly
        435                 440                 445

Lys Ser Thr Ala Thr Leu Pro Leu Ser Phe Lys Glu Ala Thr Leu Ala
    450                 455                 460

Lys Arg Phe Ala Leu Lys Thr Arg Ser Gln Ile Thr Lys Arg Lys Arg
465                 470                 475                 480

Met Ser Leu Val Lys Glu Lys Lys Ala Ala Gln Thr Leu Ser Ala Ile
                485                 490                 495

Leu Leu Ala Phe Ile Ile Thr Trp Thr Pro Tyr Asn Ile Met Val Leu
            500                 505                 510

Val Asn Thr Phe Cys Asp Ser Cys Ile Pro Lys Thr Phe Trp Asn Leu
        515                 520                 525

Gly Tyr Trp Leu Cys Tyr Ile Asn Ser Thr Val Asn Pro Val Cys Tyr
    530                 535                 540

Ala Leu Cys Asn Lys Thr Phe Arg Thr Thr Phe Lys Met Leu Leu Leu
545                 550                 555                 560

Cys Gln Cys Asp Lys Lys Lys Arg Arg Lys Gln Gln Tyr Gln Gln Arg
```

565 570 575

Gln Ser Val Ile Phe His Lys Arg Ala Pro Glu Gln Ala Leu
580 585 590

<210> 4
<211> 15
<212> DNA
<213> Homo sapiens

<400> 4
atgaacttga tgttt 15

<210> 5
<211> 15
<212> DNA
<213> Homo sapiens

<400> 5
atgaactcga tgttt 15

<210> 6
<211> 15
<212> DNA
<213> Homo sapiens

<400> 6
tcataccgtc tggca 15

<210> 7
<211> 15
<212> DNA
<213> Homo sapiens

<400> 7
tcataccatc tggca 15

<210> 8
<211> 15
<212> DNA
<213> Homo sapiens

<400> 8

ggaggagctg gggat 15

<210> 9
<211> 15
<212> DNA
<213> Homo sapiens

<400> 9
ggaggagatg gggat 15

<210> 10
<211> 15
<212> DNA
<213> Homo sapiens

<400> 10
atccagctag agtca 15

<210> 11
<211> 15
<212> DNA
<213> Homo sapiens

<400> 11
atccagccag agtca 15

<210> 12
<211> 15
<212> DNA
<213> Homo sapiens

<400> 12
aagagaatga acttg 15

<210> 13
<211> 15
<212> DNA
<213> Homo sapiens

<400> 13
gaagccaaac atcaa 15

<210> 14
<211> 15
<212> DNA
<213> Homo sapiens

<400> 14
aagagaatga actcg 15

<210> 15
<211> 15
<212> DNA
<213> Homo sapiens

<400> 15
gaagccaaac atcga 15

<210> 16
<211> 15
<212> DNA
<213> Homo sapiens

<400> 16
gggaggtcat accgt 15

<210> 17
<211> 15
<212> DNA
<213> Homo sapiens

<400> 17
accacttgcc agacg 15

<210> 18
<211> 15
<212> DNA
<213> Homo sapiens

<400> 18
gggaggtcat accat 15

<210> 19
<211> 15
<212> DNA

<213> Homo sapiens

<400> 19
accacttgcc agatg 15

<210> 20
<211> 15
<212> DNA
<213> Homo sapiens

<400> 20
gcctgaggag gagct 15

<210> 21
<211> 15
<212> DNA
<213> Homo sapiens

<400> 21
tccaccatcc ccagc 15

<210> 22
<211> 15
<212> DNA
<213> Homo sapiens

<400> 22
gcctgaggag gagat 15

<210> 23
<211> 15
<212> DNA
<213> Homo sapiens

<400> 23
tccaccatcc ccatc 15

<210> 24
<211> 15
<212> DNA
<213> Homo sapiens

<400> 24

cttcccatcc agcta 15

<210> 25
<211> 15
<212> DNA
<213> Homo sapiens

<400> 25
cacggctgac tctag 15

<210> 26
<211> 15
<212> DNA
<213> Homo sapiens

<400> 26
cttcccatcc agcca 15

<210> 27
<211> 15
<212> DNA
<213> Homo sapiens

<400> 27
cacggctgac tctgg 15

<210> 28
<211> 10
<212> DNA
<213> Homo sapiens

<400> 28
agaatgaact 10

<210> 29
<211> 10
<212> DNA
<213> Homo sapiens

<400> 29
gccaaacatc 10

<210> 30
<211> 10
<212> DNA
<213> Homo sapiens

<400> 30
aggtcatacc 10

<210> 31
<211> 10
<212> DNA
<213> Homo sapiens

<400> 31
acttgccaga 10

<210> 32
<211> 10
<212> DNA
<213> Homo sapiens

<400> 32
tgaggaggag 10

<210> 33
<211> 10
<212> DNA
<213> Homo sapiens

<400> 33
accatcccca 10

<210> 34
<211> 10
<212> DNA
<213> Homo sapiens

<400> 34
cccatccagc 10

<210> 35
<211> 10
<212> DNA

<213> Homo sapiens

<400> 35
ggctgactct 10

<210> 36
<211> 24
<212> DNA
<213> Homo sapiens

<400> 36
gcttcataga gattcagcac cctg 24

<210> 37
<211> 22
<212> DNA
<213> Homo sapiens

<400> 37
cacaattacc aggatgttgc cg 22

<210> 38
<211> 22
<212> DNA
<213> Homo sapiens

<400> 38
cgtcactcat ttcggcagct ac 22

<210> 39
<211> 23
<212> DNA
<213> Homo sapiens

<400> 39
ggaggcacag ttctctttcc aac 23

<210> 40
<211> 22
<212> DNA
<213> Homo sapiens

<400> 40

aaacgaacaa caaagagagc cg 22

<210> 41
<211> 22
<212> DNA
<213> Homo sapiens

<400> 41
tgtggtcttg gtccatctgc tc 22

<210> 42
<211> 20
<212> DNA
<213> Homo sapiens

<400> 42
gcatgaaacg ctccaacagg 20

<210> 43
<211> 25
<212> DNA
<213> Homo sapiens

<400> 43
ccgcttagtg atctgacttc tggtc 25

<210> 44
<211> 24
<212> DNA
<213> Homo sapiens

<400> 44
tggacacagc taagacttct gacg 24

<210> 45
<211> 24
<212> DNA
<213> Homo sapiens

<400> 45
gtgtgcgttt tgtcactgct attg 24